



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Posgrado en Filosofía de la Ciencia

Facultad de Filosofía y Letras

Instituto de Investigaciones Filosóficas

Filosofía de las Ciencias Cognitivas

Cognición socio-moral humana y sus implicaciones normativas: una propuesta integrativa desde el computacionalismo y el enactivismo

Tesis

Que para optar por el grado de:

Doctora en Filosofía de la Ciencia

Presenta:

Cassandra Pescador Canales

Tutor principal

Dr. Maximiliano Martínez Bohórquez, UAM-Cuajimalpa

Miembros del Comité Tutor

Dr. Eduardo García Ramírez, IIFS-UNAM

Dra. Ximena González Grandón, Facultad de Medicina-UNAM

Dra. Melina Gastelum Vargas, FFyL-UNAM

Dr. Miguel Ángel Sebastián González, IIFS-UNAM

Ciudad de México, Enero 2024



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**PROPUESTA UNIVERSITARIA DE INTEGRIDAD Y
HONESTIDAD ACADÉMICA Y PROFESIONAL
(Graduación con trabajo escrito)**

De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción 1, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la Institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado *Cognición socio-moral humana y sus implicaciones normativas: una propuesta integrativa desde el computacionalismo y el enactivismo*, que presenté para obtener el grado de **Doctorado** es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi programa de posgrado, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de graduación.

Atentamente

CASSANDRA PESCADOR CANALES
401057803

Agradecimientos

Agradezco a la Universidad Nacional Autónoma de México por brindarme la oportunidad de realizar mis estudios de doctorado. La apertura de sus puertas ha sido fundamental en mi desarrollo académico y profesional. Asimismo, agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgarme una beca nacional que respaldó financieramente mi investigación durante el periodo comprendido entre agosto de 2017 y julio de 2021.

Quiero agradecer de manera muy especial a cada uno de mis tutores, quienes creyeron en este proyecto y me ayudaron a construirlo mediante sus consejos y fructíferas aportaciones, así como también, su gran apoyo y el tiempo y la paciencia que me han brindado. Al Dr. Maximiliano Martínez Bohórquez por su invaluable ayuda, por orientarme y dirigirme pacientemente en cada paso de esta investigación y por abrirme espacios que contribuyeron significativamente a mi formación académica, en especial a PhiBio y al seminario de “Cognición y sociología de la moral” de la UAM-Cuajimalpa. Al Dr. Eduardo García Ramírez por compartir sus conocimientos y por su valiosa guía, la cual desempeñó un papel crucial en el enriquecimiento y fortalecimiento de este trabajo. A la Dra. Ximena González Grandón por su orientación invaluable para introducirme de lleno en las ciencias cognitivas corporizadas, lo que me ayudó a efectuar distinciones teóricas fundamentales. A la Dra. Melina Gastelum Vargas por sus valiosos consejos y comentarios que robustecieron significativamente mi investigación, así como también, por abrirme espacio en el seminario “cognición 4E” del IIFs, lo que ha contribuido de manera muy importante en mi formación. Y, al Dr. Miguel Ángel Sebastián por su gran apoyo y significativa instrucción.

También quiero agradecer profundamente a mi familia y amigos, quienes han tenido un papel fundamental en mi desarrollo personal y profesional, sin ustedes esto no sería posible: a mi madre por su amor incondicional y por su inagotable apoyo; a mi padre por su amor y por siempre estar; a Elidia por su solidaridad, fuerza y energía; a Alejandra Portugal por su entrañable cariño y por su apoyo en todo momento; a Jaky y a Gina por ser mi red y mi familia del corazón; a Francisco Chacón por su afectuosa compañía y apoyo y por sus acertadas ideas que me ayudaron a encontrar el camino en la investigación. Y, a Iván, a Yael y a Axel por sus grandes enseñanzas.

Por último, también agradezco a mis compañeros y compañeras del posgrado, con quienes compartí espacios llenos de fructíferas discusiones y motivaciones, lo que me enriqueció enormemente y me impulsó a continuar. Agradezco la colaboración y el valioso intercambio de ideas que hemos compartido, lo que contribuyó significativamente a mi desarrollo durante este periodo.

¡A todos ustedes mi reconocimiento y gratitud!

A Dios
A María Elidia Canales Ramos
A Ricardo Pescador Elizondo

Índice

I)	Introducción	7
Capítulo 1. Antecedentes: cartografía del debate		
1.1.	Introducción	12
1.2.	Evolución: Escuelas de la psicología evolucionista	12
1.3.	Cognición: fundamentos y supuestos	17
1.3.1.	Computacionalismo	17
1.3.2.	Enactivismo	22
1.3.3.	Consecuencias	26
1.4.	Cognición social	27
1.4.1.	Atribucionistas	28
1.4.2.	Cognición social corporizada	36
1.4.3.	Consecuencias	43
1.5.	Cognición moral	44
1.5.1.	Psicología del desarrollo moral	44
1.5.2.	Aproximaciones enactivas a la moral	48
1.5.3.	Consecuencias	50
II)	Preámbulo a los capítulos 2 y 3	51
Capítulo 2. Computacionalismo		
2.1.	Introducción	56
2.2.	Evolución	59
2.2.1.	Problema de origen	59
2.2.2.	Psicología evolutiva: escuela de Santa Bárbara	62
2.2.3.	Desarrollo del argumento evolutivo	63
2.2.4.	Consecuencias	70
2.3.	Cognición: supuestos cognitivos computacionalistas	70
2.3.1.	Procesamiento cognitivo- funcionalista	71
2.4.	Cognición social computacionalista	72
2.4.1.	Cognición social modularista	74
2.4.2.	Consecuencias	81
2.5.	Cognición moral computacionalista	81
2.5.1.	Justificaciones	82
2.5.2.	Arquitectura súper-modular: superando las restricciones de los viejos debates	90
2.5.3.	Moral súper-modular	92
2.5.4.	Problemas con el modularismo tradicional	95
2.5.5.	Argumento del NMI	98
2.5.6.	¿De qué se constituye el NMI?	100
2.5.7.	Evidencia empírica	101
2.5.8.	Algunas anotaciones respecto al desarrollo posterior de la moral	111
2.5.9.	Línea cronológica del NMI: el desarrollo moral pre-lingüístico	114
2.6.	Normatividad	115
2.6.1.	Propuesta normativa: conocimiento moral general objetivo	115
2.6.2.	Argumento normativo condensado	119
2.6.3.	Consecuencias	120
Capítulo 3. Enactivismo		
3.1.	Introducción	121
3.2.	Evolución	122
3.2.1.	Enactivismo y construcción de nicho	122
3.2.2.	Construcción de Nicho	123

3.2.3.	Aportaciones directas de la hipótesis de la construcción de nicho a la cognición socio-moral humana	126
3.2.4.	Antecedentes de la intersubjetividad: evolución del uso de símbolos	127
3.2.5.	Transición de la mimesis corporal al lenguaje	131
3.2.6.	Consecuencias	136
3.3.	Cognición enactiva: modelo cognitivo	136
3.3.1.	Consecuencias	138
3.4.	Cognición social enactiva	138
3.4.1.	Interacción socio-cognitiva	139
3.4.2.	Normatividades enactivas	140
3.4.3.	Consecuencias	145
3.5.	Cognición moral enactiva	145
3.5.1.	Normatividad moral enactiva	145
3.5.2.	Normatividad fenomenológica	149
3.5.3.	Normatividad moral socio-cultural	151
3.5.4.	Normatividad moral enactiva: la propuesta condensada	155
3.5.5.	Consecuencias	158
3.6.	Facultades socio-morales: intervenciones corporizadas	159
3.6.1.	Implicaciones problemáticas de la cognición social computacionalista	160
3.6.2.	Argumento	163
3.6.3.	Evidencia empírica	166
3.6.4.	Evidencia para facultades morales tempranas: conducta cooperativa mínima	176
3.6.5.	Consecuencias	182

Capítulo 4. Hacia una propuesta integrativa computacionalista-enactiva de la cognición socio-moral y sus implicaciones normativas

4.1.	Introducción	184
4.1.1.	Condiciones mínimas para que sea un caso computacionalista o un caso enactivo	187
4.1.2.	Compromisos asumidos que habilitan la propuesta integrativa computacionalista - enactiva.	188
4.2.	Arquitectura cognitiva PI	188
4.2.1.	¿Qué es lo que habilita a la arquitectura cognitiva PI?	190
4.2.2.	Arquitectura PI como metáfora cognitiva	196
4.2.3.	Distinción entre PI y la arquitectura dual socio-cognitiva	198
4.2.4.	Justificación de la pertinencia de PI	198
4.3.	Evolución	200
4.3.1.	Convergencia bio-cultural	202
4.3.2.	El cuerpo vívido como punto de inflexión entre lo biológico y lo cultural	205
4.4.	Cognición computacionalista-enactiva	210
4.5.	Cognición social computacionalista-enactiva	212
4.6.	Cognición moral computacionalista-enactiva	213
4.7.	Algunos problemas que pueden surgir	215
4.8.	Normatividad moral computacionalista-enactiva	217
4.8.1.	Esquema del modelo cognitivo de la normatividad computacionalista-enactiva	220
4.9.	Co-participación de los elementos presentados	221
4.10.	Respuestas a las dificultades presentadas: articulando argumentos	228
4.10.1.	Evolución	228
4.10.2.	Cognición	229
4.10.3.	Cognición social	230
4.10.4.	Cognición moral	231
4.10.5.	Computacionalismo y enactivismo en tareas socio-morales	233
4.11.	Normatividad	233
	5. Conclusiones	236
	6. Bibliografía	239

Introducción

En la presente investigación doctoral me centro en el estudio de la cognición socio-moral humana con el objetivo de definir los criterios teóricos y empíricos que den pie a un modelo explicativo integral de la misma en el marco de la filosofía de las ciencias cognitivas. Para ello, empleo una metodología pluralista epistémica que permita argumentar a favor de un modelo flexible, posibilitando así, caracterizar la naturaleza de los procesamientos y elementos que le constituyen, así como también, explicar sus implicaciones normativas y filosóficas.

El trabajo se dedica a explorar a profundidad las aportaciones teóricas-empíricas provenientes de dos marcos conceptuales contrastantes de la ciencia cognitiva: el “funcionalismo” en su versión concerniente a “la teoría computacional de la mente” (o “computacionalismo”), por un lado, y las “ciencias cognitivas corporizadas”, por el otro, en donde se sitúa el “enactivismo”, cuyo programa de investigación es en el que me baso específicamente. La comparación de ambos marcos conceptuales permite confrontar de manera general dos diferentes modos y estadios de la disciplina, pero, a su vez, este ejercicio permite detectar los elementos que constituyen una explicación satisfactoria del fenómeno a estudiar.

En concreto, mi argumento es que tanto el “computacionalismo” como el “enactivismo” otorgan aportaciones significativas y ofrecen contribuciones correctas, distintivas e importantes a las explicaciones de la cognición socio-moral humana, pero lo hacen enfocándose en distintos aspectos de la misma, por lo que propongo, a partir de un pluralismo pragmático, un marco integrativo capaz de considerar y abordar lo valioso de cada uno de ellos que, además, permita escudriñar sus implicaciones normativas.

Así, la presente tesis tiene como objetivo abordar dos temas centrales a partir de los postulados de ambos marcos conceptuales:

1. Lo cognitivo: enfocado en explicar los procesamientos cognitivos involucrados en la competencia socio-moral, tanto los representacionales guiados por principios como aquellos habilitados por acoplamientos sensorio-motores beneficiados por bucles de percepción-acción.
2. Lo normativo: enfocado en ofrecer una herramienta evaluativa moral capaz de generar juicios *situados* pero anclados a criterios *objetivos*.

En este punto es importante aclarar la razón por la que le llamo “cognición socio-moral” al fenómeno cognitivo a estudiar. La razón es porque se ha considerado que la moral es una sofisticación de nuestras competencias sociales y, como tal, se ha mostrado que las concepciones de la cognición social tienen implicaciones directas en cómo se estudia la cognición moral. De ello, se ha concebido que la cognición social es una condición necesaria de la cognición moral, lo que resulta en un constructo conceptual claro que efectivamente se le puede llamar “socio-moral”. En esta línea, el énfasis de estudio en la presente tesis es en el aspecto moral, pero dado que éste siempre es atravesado por las concepciones socio-cognitivas que se asumen, en adelante usaré en sinonimia “cognición moral” y “cognición socio-moral”.

Ahora bien, como es de esperarse, el computacionalismo y el enactivismo representan paradigmas distintos, por lo que existen claras divergencias teóricas y metodológicas entre ellos, las cuales pueden verse como una virtud y no como un problema. Pero, si esto es así, entonces, el primer desafío a superar es ¿existe alguna arquitectura cognitiva capaz de soportar los dos modos de cognición (computacionalista y enactivo) al mismo tiempo que respete los principios y supuestos que cada uno sostiene?

Frente a esta pregunta mi hipótesis es que sí, y para sustentarla, propongo explorar los nuevos desarrollos que han surgido en el campo de la inteligencia artificial (IA), donde se verá que en la actualidad existen modelos llamados *programas informáticos de inteligencia artificial* (PI en adelante) capaces de soportar al menos dos tipos de arquitecturas: i) una de base computacional favorecida por el uso de principios, representaciones y ejecutada por procesamientos de información, y; ii) otra arquitectura habilitada por un “lenguaje de programación avanzada” conocido como *Deep Learning* (DL en adelante), el cual es capaz de soportar que parte de su diseño sea *aprendido* en interacción con el entorno y mediante la *experiencia*, donde su concepto central es la *acción*.

De ahí, mi propuesta es que el computacionalismo se correspondería con (i), mientras que el enactivismo se correspondería con (ii), con lo que pretendo defender que es a partir de esta posible metáfora como podemos cumplir el doble compromiso metodológico-cognitivo a la vez que se respetan los principios de cada uno, al menos, para la cognición socio-moral.

Considero que PI puede servir como una posible metáfora de arquitectura cognitiva por ser capaz de aceptar ambos modos de cognición, de lo que planteo, derivan beneficios y ventajas metodológicas.

Si mis consideraciones son correctas, entonces es plausible concebir que los dos modos de cognición (computacionalista y enactivo) son constitutivos de nuestra cognición socio-moral, cada uno respondiendo a distintos aspectos de la misma: el computacionalismo enfocado en los procesamientos de información guiados por principios y el enactivismo enfocado en el aspecto *relacional* que emerge de las interacciones agente-entorno, con lo que busca capturar sus dimensiones fenomenológicas, interactivas y corporizadas. Bajo esta lectura, es posible proponer la siguiente hipótesis:

Hipótesis: La cognición socio-moral humana se constituye de procesamientos computacionalistas y enactivos, de tal manera que:

El computacionalismo:

- 1) En lo cognitivo: explica los procesamientos representacionales guiados por principios involucrados en las facultades morales, tales como: *teoría de la mente*, entendimiento intencional, entendimiento emocional, principios morales pre-lingüísticos y empatía, los cuales son procesamientos de información cargados de contenido que se benefician de las interacciones que tienen entre sí.
- 2) En lo normativo: explica la presencia de un tipo de conocimiento moral general dotado evolutivamente que es capaz de asegurar la objetividad moral.

El enactivismo:

- 1) En lo cognitivo: explica los aspectos corporizados y el papel que desempeñan los procesos interactivos en la cognición socio-moral, lo que se logra por otorgar autonomía a los procesos supra-personales.
- 2) En lo normativo: explica la normatividad moral *situada* que deriva de su metodología *relacional*.

Para lograr mostrarlo, parto de una pregunta clave que guiará la investigación: ¿qué explica y qué resuelve empíricamente una metodología que no explique y no resuelva la otra?

Con el objetivo de responder a ello, elaboro un análisis detallado de los fundamentos que habilitan a cada marco teórico para distinguir claramente las contribuciones que cada uno

puede ofrecer, para lo cual es preciso analizar tres objetos de estudio que están detrás de ellos: 1) sus compromisos evolutivos (respecto a la psicología evolucionista), 2) el procesamiento cognitivo que postulan y los supuestos cognitivos que sostienen, y, 3) sus explicaciones acerca de la cognición social. Una vez explicados estos tres puntos, es posible abordar al cuarto tema: 4) sus explicaciones y aportaciones a la cognición moral, así como también, sus implicaciones normativas.

De ahí, entonces, la estructura de la tesis es de la siguiente manera:

En el capítulo 1, llamado “Antecedentes”, presento la cartografía de los debates que atraviesan la discusión, es decir, presento un panorama claro de los antecedentes que constituyen los debates relevantes precedentes de cada marco conceptual. El capítulo 2 se dedica a explorar el marco computacionalista; el capítulo 3 se ocupa de explicar el marco enactivo y en el capítulo 4 presento mi propuesta integrativa.

Cada capítulo lo divido en los siguientes cuatro temas, cada uno conformando una subsección: a) Evolución (desde la psicología evolucionista), b) Cognición: procesamiento cognitivo, c) Cognición social y d) Cognición moral.

Considero que de ello se verá que el computacionalismo y el enactivismo pueden complementarse coherentemente, y propongo que esto es posible si se acepta dicha arquitectura cognitiva PI.

Por último, la justificación de mi propuesta integrativa se basa al menos en cinco razones: 1) debido a que cada marco conceptual se enfoca en estudiar distintos aspectos de la cognición socio-moral, ambos son capaces de ofrecer distintas explicaciones, por lo que su abordamiento posibilita un análisis más abarcador que pretende arribar a la integración de sus elementos; tal explicación integracionista involucraría propiedades distintas a fin de satisfacer la red de objetivos asociados al *explanandum*; 2) Francisco Varela (1999), entre otros autores, ha reconocido la distinción entre los juicios y las justificaciones morales por un lado y el comportamiento moral (inmediato) por el otro, sosteniendo que el segundo tema ha recibido menor atención en las agendas disciplinares, por lo que parece necesario ofrecer explicaciones del mismo. Siguiendo a Varela, considero que abordar una perspectiva *relacional* basada en el análisis de la *acción* que emerge del confrontamiento directo con el mundo, puede contribuir significativamente a ello; 3) los recientes desarrollos conceptuales y empíricos de la cognición social han señalado la necesidad de sumar a las

teorías disponibles (tales como las computacionalistas) investigaciones que estudien los encuentros sociales en tiempo real de una manera verdaderamente *interactiva* (Schilbach, Timmermans, Reddy, Costall, Bente, Schlicht, y Vogeley, 2013), es decir, de una manera *relacional*; exigencia que el enactivismo es capaz de satisfacer; 4) discursos sociológicos de la moral plantean estudiarla en contextos pragmáticos *relacionales* que la conciban como una estructura social. Bajo esta concepción, las estructuras sociales son almacenadas y procesadas en sujetos individuales, pero también se componen de patrones duraderos de significado que emergen socialmente y como tal adquieren independencia por trascender el nivel individual; así, se ha subrayado la importancia del aspecto *relacional* de la moral, cuyo análisis debe ganar un lugar en el estudio de la cognición socio-moral, y; 5) respecto a la reconocida brecha entre la filosofía analítica y la filosofía continental, algunos filósofos morales han buscado un posible acercamiento entre ellas (p. ej. Macintyre, 1981; Williams, 2000; Foot, 2002), considerando que habría una gran ganancia en su reconciliación (ver Edwards, 2017). Este posible acercamiento invita a cruzar las fronteras entre marcos conceptuales contrastantes y con ello realizar un escrutinio profundo que permita un diálogo iluminador, a fin de advertir las virtudes que cada uno puede ofrecer.

En este contexto, busco proponer que tanto los procesamientos de información guiados por principios (computacionalistas) como los procesos de *hacer-sentido* (enactivos) tienen un papel habilitador en la cognición socio-moral, y frente a ello, me he trazado la tarea de demarcar cuál es el alcance e implicación de cada uno. De esta manera, mi objetivo es proponer la posible integración de ambos modos-cognitivos y con ello destacar las aportaciones explicativas que cada uno ofrece al complejo fenómeno de la cognición socio-moral humana.

Capítulo 1

Antecedentes: cartografía del debate

1.1. Introducción

En este capítulo ofrezco una visión panorámica respecto a los debates que han tenido lugar en relación con las explicaciones de la psicología evolutiva, de los supuestos cognitivos, socio-cognitivos y de la cognición moral pertinentes al computacionalismo y al enactivismo, respectivamente. Así, este capítulo está dedicado a mostrar la cartografía de los debates relevantes que preceden y atraviesan a cada marco conceptual. Para ello, la estructura de esta sección la divido en cuatro apartados: 1) evolución, 2) cognición, 3) cognición social y 4) cognición moral.

En la sección de “Evolución”, me ciño a presentar las aproximaciones de las escuelas de la psicología evolucionista, en particular, hago explícito de dónde parten y qué intenta explicar cada una de ellas¹. En la sección de “Cognición”, abordo los modelos, supuestos y procesamientos cognitivos que están detrás de cada marco conceptual, es decir, explico brevemente de dónde proviene cada uno de ellos y los compromisos metodológicos, ontológicos y epistemológicos que los fundamenta. En la sección de “Cognición social”, expongo el debate que prevalece entre las explicaciones computacionalistas y enactivas pertinentes a nuestras competencias socio-cognitivas, lo que resulta en implicaciones fundamentales para la cognición moral. Y, por último, en la sección de “Cognición moral”, presento los debates y las aportaciones provenientes de la psicología del desarrollo y de la neurociencia que han significado el progreso del cognitivismo clásico respecto al tema de interés, asimismo, presento un breve panorama de las aportaciones enactivas que se han hecho hasta el momento respecto a la moral.

1.2. Evolución: escuelas de la psicología evolucionista

¹ Es preciso mencionar que el enactivismo en general se aleja de la carga conceptual de los supuestos evolucionistas tradicionales, esto por defender una evolución del organismo *in situ*. No obstante, se verá que es plausible relacionar al enactivismo con la escuela de la “construcción de nicho” debido a que ésta hace justicia a sus principios teóricos, tema que desarrollo en el capítulo 3.

Abordo las escuelas de la psicología evolucionista debido a la importancia e impacto directo que han tenido respecto a los estudios y a los avances en las ciencias cognitivas, partiendo desde su mayor precedente que es la sociobiología.

En la década de los 70 surge la sociobiología como un programa de investigación que intenta explicar la conducta social a partir de argumentos biológicos innatistas. De ahí, en los 80 nace la psicología evolutiva como un descendiente directo de la sociobiología con una clara diferencia: “La sociobiología se centra en el *comportamiento* innato, mientras que la psicología evolutiva se centra en los *mecanismos psicológicos* subyacentes a ese comportamiento” (Joyce, 2006, p. 5, las cursivas son mías).

En este contexto, se ha creído que sólo existe una psicología evolutiva que intenta explicar tales mecanismos exclusivamente desde el reconocimiento de una sola fuerza evolutiva: la *selección natural*. No obstante, en un análisis más detallado de su literatura, algunos académicos han destacado y clasificado tres escuelas principales que si bien, comparten un mismo supuesto (que la mente humana ha atravesado un proceso de evolución), también presentan serias diferencias teóricas y metodológicas entre ellas:

- 1) La Escuela de Santa Bárbara o de la mente modular
- 2) La Escuela Británica o del cerebro social
- 3) La Escuela Cultural o del cerebro cultural

A continuación, presento sucintamente los planteamientos que constituyen a cada una de ellas.

Respecto a la “Escuela de Santa Bárbara” (1992), ésta se consolida como la iniciadora de la psicología evolutiva y es fundada por la psicóloga Leda Cosmides y por los antropólogos Jerome Barkow y John Tooby, quienes pertenecían a la Universidad de California de Santa Bárbara, de ahí su nombre. Uno de sus grandes representantes (aunque no es perteneciente de la escuela) es Steven Pinker, quien ha difundido ampliamente sus propuestas.

Los supuestos evolutivos defendidos por la “Escuela de Santa Bárbara” son principalmente el *modularismo*, el *adaptacionismo* y el *aculturalismo*, lo que resulta en la *hipótesis de la mente modular*. Así, su tesis central es que existe una naturaleza humana universal a nivel de mecanismos psicológicos evolucionados. Esta tesis es desarrollada en su obra principal titulada *The adapted mind* (1992). En breve:

“La premisa central de *The adapted mind* es que existe una naturaleza humana universal, pero que esta universalidad existe principalmente a nivel de mecanismos psicológicos evolucionados, no en conductas culturales expresas. Para esta perspectiva, la variabilidad cultural no es un reto a los llamados de universalidad, sino una data que puede darnos alguna revelación sobre la estructura de los mecanismos psicológicos que ayudaron a generarla. Una segunda premisa es que estos mecanismos psicológicos evolucionados son adaptaciones construidas por la selección natural a lo largo del tiempo. Una tercera premisa [...] es que la estructura evolucionada de la mente humana está adaptada a la forma de vida de los cazadores-recolectores del Pleistoceno y no necesariamente a nuestras circunstancias actuales” (Barkow, Cosmides y Tooby, 1992, p. 5).

En suma, el *adaptacionismo* refiere a que ciertos rasgos psicológicos son el resultado de adaptaciones al medio moldeadas por selección natural y transmitidas genéticamente, de ahí que a los diversos comportamientos se les considere como “adaptaciones”. El *modularismo*, derivado del *adaptacionismo*, sostiene que la mente humana cuenta con módulos innatos especializados para la resolución de diversos problemas moldeados por la selección natural (ej. existe un módulo para el lenguaje, un módulo para la percepción de emociones, un módulo para el reconocimiento de rostros etc.). Y, por último, se le acusa de *aculturalista* debido a que suele relegar a la cultura de sus explicaciones y, en su lugar, prima la “selección natural” como mecanismo explicativo de las adaptaciones complejas.

Por su parte, la “Escuela Británica” (1998) surge para contrarrestar los postulados de la “Escuela de Santa Bárbara” y se presenta como una alternativa que vincula la evolución humana y la cultura. Esta escuela es fundada por Robin Dunbar y seguida por Louise Barrett y John Lycett, entre otros. Su postulado principal es la *hipótesis del cerebro social*, de acuerdo con la cual, la mente humana evolucionó por nuestra “intensa vida social y cooperativa” expresada en nuestros complejos sistemas sociales, es decir, la mente humana evolucionó principalmente al lidiar con grupos sociales complejos y con fenómenos tales como la caza, el lenguaje, la religión, etc.

En particular, Dunbar muestra que el acicalamiento (*grooming*) físico es un mecanismo importante de unión social y favorece la formación de coaliciones en primates. Dunbar basa su teoría en correlaciones sólidas entre el tamaño del grupo y el tamaño del neocórtex en primates, desde monos hasta humanos. La propuesta central es que el neocórtex aumentó en respuesta a pertenecer a grupos más grandes: "Dado que mantener grupos coherentes es cognitivamente exigente, el tamaño del cerebro (o más específico, el volumen del neocórtex) evolucionará para satisfacer las demandas cognitivas del tamaño del grupo óptimo de la especie". (Sutcliffe *et al.*, 2012, p. 51, en Etzelmüller, 2016, p.159).

Así, la Escuela Británica acusa a la Escuela de Santa Bárbara de ser “rígidamente racional” (Gamble *et al.*, 2014) y de derivar en cierto determinismo genético. En su lugar, Dunbar se coloca en el otro extremo y se orienta a una visión que relaciona la evolución humana con la cultura y el entorno.

En esta línea, la tercera escuela de la psicología evolutiva, conocida como la “Escuela Cultural” (2005, 2010, 2012), también surge de cuestionar y desafiar los postulados *modularistas* y *adaptacionistas* propios de la Escuela de Santa Bárbara y, en su lugar, postula la *hipótesis del cerebro cultural*.

La “Escuela Cultural” se puede clasificar en tres visiones: 1) la hipótesis del cerebro cultural, 2) la teoría de la herencia dual (conocida también como la co-evolución gen-cultura) y, 3) la teoría de la construcción de nicho (actualmente se ramifica en lo que se conoce como “genómica cultural”, que busca corroborar empíricamente los supuestos que sostiene).

La relación de estas tres versiones es que *la hipótesis del cerebro cultural* (Henrich, 2011; Jaquet, 2012) proviene directamente de *la teoría de la herencia dual* (Richerson y Boyd 1985, 2005) y se estrecha contundentemente con *la teoría de la construcción de nicho* (Laland, Odling-Smee, y Myles, 2010), siendo los principales representantes de esta última Kevin Laland, Michael Tomasello, Cecilia Heyes y Joseph Henrich, entre otros.

El principal planteamiento de la Escuela Cultural destaca por relacionar potencialmente la evolución con la “influencia cultural”, dando así un papel preponderante e irrevocable a la cultura en sus explicaciones evolucionistas, y, en consecuencia, reconoce el carácter *activo* de la evolución humana. En otras palabras, desde este enfoque, el ser humano se adapta a diversos entornos por modificarlos, lo que logra debido a sus altas capacidades sociales y culturales y, de esta manera, se le otorga un espacio relevante al *aprendizaje cultural*.

En este contexto, la visión de la *construcción de nicho* defiende que: “Los seres humanos no modifican su entorno al azar. Más bien, al igual que otros animales que construyen nidos, montículos, redes y presas que los ayudan a ellos y a su descendencia a sobrevivir, los humanos construyen estructuras y tienen otros impactos en su mundo que mejoran ampliamente su aptitud evolutiva” (Laland, 2017, p. 231).

En particular, la *hipótesis del cerebro cultural* se fundamenta en *la teoría de la herencia dual* (o teoría de la co-evolución gen-cultura) postulada por el antropólogo Robert Boyd y el biólogo Peter Richerson (1985), la cual defiende que el cerebro humano co-evolucionó con

las estrategias de aprendizaje y el tamaño del grupo social, entre otros factores; es decir, apuestan por la co-evolución sincrónica entre la cultura y los genes hasta constituir la como la primera fuerza evolutiva que dio lugar a la evolución humana.

Los teóricos que se ciñen a la *hipótesis del cerebro cultural* enfatizan “cómo el cambio cultural puede limitar o acorralar el cambio genético, tanto, como las psicologías de base genética influyen en el curso del cambio cultural” (Richerson y Boyd, 2006, p. 194).

De manera importante, en la actualidad, tales planteamientos son corroborados por el enfoque que se conoce como “genómica cultural” (2018), que defiende la *interacción* entre la genética y la cultura y estudia su *influencia recíproca*, con lo que abandona la manera tradicional de estudiar a esta dupla, i.e. ya no estudia en *paralelo* las interacciones gen-cultura. Esto ha significado un enorme progreso para las posturas que siguen la *hipótesis del cerebro cultural*.

Una cuestión relevante es que en la línea de estudio de la genómica cultural, ya no se considera el supuesto tradicional sostenido por algunas explicaciones evolutivas, a saber, que las diferencias entre *grupos* humanos se deben 100% a variaciones culturales y no a variaciones genéticas (pues se ha sostenido que las similitudes genéticas entre los grupos humanos son del 99.9%, lo que deja un porcentaje muy pequeño para la variación, y ésta sólo es pensada en las diferencias de atributos superficiales como el color de piel, la pigmentación de ojos, cabello, etc., que comprendería ese 0.01% restante). Al respecto, los teóricos de la genómica cultural sostienen que esto sólo es posible a nivel teórico, pues gracias a las nuevas tecnologías se ha descubierto que sí hay diferencias importantes entre los grupos humanos a nivel biológico, con más de las tres cuartas partes de las variantes de ADN que difieren entre los grupos, mismas que se corresponderían con las diferencias culturales; la mayoría de origen evolutivo reciente (entre 5.000 y 10.000 años) (Chen y Moyzis, 2018, p.765).

Bajo estos argumentos se ha desafiado a la “teoría neutra” (Kimura, 1983) que había prevalecido hasta hace muy poco tiempo, la cual sostiene que las variaciones entre grupos humanos (más allá de las que comprenden ese 0.01%) son aleatorias e intrascendentes. Así, el principal supuesto de la genómica cultural que desafía empíricamente a las concepciones tradicionales es que “nuestra cultura es nuestro entorno y las diferencias culturales pueden dirigir a cambios en nuestro genoma” (Chen y Moyzis, 2018, p.765), con

lo que defienden la fuerte *influencia recíproca* que existe entre la genética y la cultura y, con ello, se posicionan como un poderoso enfoque para los estudios evolucionistas actuales.

Es valioso considerar estas explicaciones, pues el número y tipo de mecanismos que favorecen la evolución y los niveles en los que puede ocurrir, ha proporcionado nuevos y fructíferos modelos para comprender a los organismos y sus interacciones. (Joyce, 2018, p. 296).

Para mis propósitos, es suficiente este breve esbozo de las escuelas de la psicología evolucionista, pues revela los datos relevantes que atraviesan los supuestos de los que parte cada marco conceptual, por lo que sus detalles se verán en la sección pertinente al marco que los sostiene: el computacionalismo se circunscribe plausiblemente a la Escuela de Santa Bárbara y el enactivismo a la Escuela Cultural, en particular, a la visión de la construcción de nicho y a la evidencia empírica de la genómica cultural.

Ahora, corresponde abordar los fundamentos y supuestos cognitivos que están detrás del computacionalismo y del enactivismo, respectivamente.

1.3. Cognición: fundamentos y supuestos

1.3.1. Computacionalismo

En el marco de la filosofía de la mente el “computacionalismo” es una versión prominente del “funcionalismo”, el cual defiende que “estar en un estado mental es estar en un estado funcional”. Pero, ¿qué es un estado funcional? éste se explica por sus relaciones causales, es decir, debe tener una causa y un efecto, como una entrada (*input*) y una salida (*output*), siendo los estados mentales a la vez causas y efectos; una causa individualizada, específica y concreta que resulta en una cierta salida. En otras palabras, los estados mentales son una correlación de causas o entradas sensoriales y efectos conductuales o efectos materializados físicamente. Así, el funcionalismo parte de la pregunta ¿para qué es la mente? y no de la cuestión ¿qué es la mente? lo que trata de resolver desde su aspecto funcional, o bien, desde un enfoque de correlaciones (contingentes) causa-efecto (Putnam, 1960; Lewis, 1980; Fodor, 1984).

En breve, el funcionalismo es la tesis filosófica que afirma que los estados mentales son esencialmente definidos por el conjunto de relaciones causales (contingentes) entre: a) estímulos del entorno; b) estados mentales y; c) la conducta (p. ej. Churchland, 1992). En consecuencia, en el marco funcionalista los estados mentales son considerados como

estados que ocupan un rol causal dentro de una red causal, entonces, los estados mentales son entendidos como causalmente conectados a otros estados mentales, al entorno y a la conducta. Así, pues, los estados mentales son estados funcionales y como tal pueden realizarse de múltiples maneras.

En estos supuestos funcionalistas se basa el computacionalismo, por tanto, aquí se está hablando del funcionalismo-computacionalista, versión predominante en las ciencias cognitivas (pues existen otras versiones del funcionalismo tales como el funcionalismo conexionista, el funcionalismo extendido y el funcionalismo mecanicista).

Ahora bien, el "Computacionalismo" o "Teoría computacional de la mente" (Hilary Putnam, 1961, 1967; Jerry Fodor, 1975, 1981, 1987, 1990, 1994, 2008) se basa en la metáfora computacional, en la que el cerebro se concibe como un ordenador y así se plantea que la mente es básicamente un sistema de procesamientos de información intracraneal que manipula representaciones simbólicas y, por tanto, concibe la cognición como un proceso computacional que calcula (o *computa*) la información mediante un determinado algoritmo, enfatizando así procesamientos que implican representaciones mentales internas y que siguen determinadas reglas. La naturaleza computacional de la cognición resalta la importancia de algún tipo de estructura de representación interna.

Siguiendo los principios funcionalistas, la motivación de la metáfora computacional es que el computador clásico toma símbolos estructurados "similares al lenguaje" como entrada (*input*) y produce símbolos estructurados "similares al lenguaje" como salida (*output*). El perfil entrada-salida (*input-output*) es mediado por la aplicación de reglas que en conjunto constituyen el programa que la máquina ejecuta. Las reglas que manipulan los símbolos es un proceso mecánico que no requiere que la máquina entienda o aprecie el significado de los símbolos que está manipulando. La máquina es sensible a las propiedades sintácticas o formales de los símbolos y las reglas que manipulan los símbolos se aplican en relación con tales propiedades. Los símbolos manipulados comúnmente tienen propiedades semánticas. Además, el significado de los símbolos de salida (*output*) generados por una computadora clásica está coherentemente relacionado con el significado de los símbolos de entrada (*input*) (ver Caín, 2016).

Así, el perfil entrada-salida refiere a que el sujeto cognitivo percibe la información proveniente del entorno, la procesa (mediante representaciones y funciones algorítmicas que siguen ciertas reglas) y el resultado es una decisión que luego se convierte en acción,

aunque, eso sí, un estado mental determinado puede desembocar en diferentes realizaciones materiales, es decir, posee *realizabilidad múltiple*² en el sentido de que una clase mental puede ser realizable por clases físicas completamente distintas.

Ahora bien, detrás de estas ideas hay compromisos y supuestos filosóficos que deben ser explicitados, los cuales abordo a continuación.

- **Supuestos computacionalistas**

En particular, el funcionalismo computacionalista es propuesto por Hilary Putnam en 1961 y desarrollado por Jerry Fodor (principalmente en los años 60 y 70), y se basa en la fusión del "realismo intencional" y del "fiscalismo".

El "fiscalismo" es una doctrina de corte monista materialista y defiende que sólo existe una sustancia: lo físico (i.e. el cerebro), lo que resulta en suponer la inexistencia de lo mental como otra sustancia ontológicamente distinta a lo físico, es decir, niegan que lo mental sea distinto de lo físico: no hay dos sustancias, por lo que hay cierta reducción de lo mental a lo físico.

El "realismo intencional", por su parte, pone en primer plano la *causalidad* de los estados mentales y las *actitudes proposicionales* (*creer que, desear que, pensar que, etc.*). En breve, se denomina "actitud proposicional" al estado mediante el cual se relaciona un sujeto con el contenido de un estado mental, por ejemplo, *Juan cree que Londres es una ciudad sofisticada*. Estos estados mentales son intencionales debido a que poseen lo que F. Brentano consideró "intencionalidad", es decir, la propiedad de "dirigirse a un contenido" y "relacionarse con un contenido". Así, la intencionalidad involucra un tipo particular de actitud hacia aquel contenido (Brunsteins, 2010, p. 30).

Sostener que tales estados poseen contenido significa que tratan sobre lo que las personas creen, desean, piensan, etc. acerca de algún rasgo del mundo y, por tanto, son estados mentales que poseen condiciones de satisfacción en correspondencia con el mundo, es

² El argumento de la "Realizabilidad Múltiple" se da principalmente en torno al problema mente-cuerpo en el marco de la filosofía de la mente, argumento que permite dar el paso de la *Teoría de la Identidad* al *Funcionalismo*. En breve, la Teoría de la Identidad sostiene que los estados mentales son idénticos a los estados cerebrales, tesis que socava al *Conductismo*. Hilary Putnam en "Psychological Predicats" (1967) propone el argumento de la Realizabilidad Múltiple con el fin de desafiar lo postulado por la *Teoría de la Identidad*. Fodor en "Special Sciences" (1974) generaliza el argumento de la Realizabilidad Múltiple y plantea que los estados mentales no son, en principio, reducibles a propiedades físicas, refutando así al reduccionismo radical.

decir, son semánticamente evaluables, pues su contenido relaciona al estado mental con una parte del mundo. Esto es relevante debido a que la *causación* entre los estados mentales, el entorno y la conducta se da en virtud de su contenido³, así como también, tiene que ver con la actitud o intención que tiene el sujeto hacia la proposición que expresa el contenido de tal estado mental.

De esta manera, el “realismo intencional” es una tesis que afirma la existencia de la causación mental y aceptar la causalidad en la esfera de lo mental implica aceptar la existencia de principios o generalizaciones del sentido común que sostienen esas relaciones causales. Fodor propone una definición de lo que es ser *realista* acerca de las actitudes proposicionales:

"Propongo decir que alguien es un Realista acerca de las actitudes proposicionales si y sólo si (a) sostiene que hay estados mentales cuyas ocurrencias e interacciones causan la conducta, y lo hacen, además, de manera que (al menos aproximadamente) respeten las generalizaciones de la psicología del sentido común acerca de los deseos y las creencias; y (b) sostiene que esos mismos estados mentales causalmente eficaces son también semánticamente evaluables." (Fodor, 1991).

Así, Fodor sostiene que los estados mentales (como actitudes proposicionales) son entidades que satisfacen las siguientes características: a) son semánticamente evaluables; b) poseen poderes causales y; c) sus generalizaciones implícitas son verdaderas.

Respecto a la evaluación semántica, se refiere al hecho de que las actitudes proposicionales poseen contenido y tal contenido tiene condiciones de satisfacción respecto a su relación con el mundo. Ahora bien, el poder causal de los estados mentales también se da en función de su contenido, pues éste afecta la manera en que las actitudes proposicionales se relacionan con otros estados mentales o con el entorno. De ello tenemos que, desde este enfoque, la propiedad semántica de las actitudes está fuertemente ligada a sus poderes causales, es decir, las actitudes proposicionales se asumen como *semánticamente evaluables* y *causalmente eficaces*, lo que quiere decir que los estados mentales, en virtud de su contenido, son concebidos como causantes de la conducta y, a su vez, los estados mentales son causados por los estímulos del entorno o por otros estados mentales.

³ Es pertinente mencionar que, según algunas teorías, el rol causal no basta para fijar el contenido. No obstante, en la noción fodoriana el rol causal es preeminente, esto debido a que se asume la tesis del *Realismo Intencional*.

Conocer el contenido de los estados intencionales es relevante porque permite individualarlos, permite distinguirlos a unos de otros, cuya funcionalidad radica en que podemos diferenciar, tanto explicativamente como predictivamente, cuál estado mental causa la conducta, en otros términos, desde esta concepción, los estados mentales son funcionalmente discretos, puesto que son entidades que pueden discriminarse efectivamente unas de otras.

Así, pues, en suma, el funcionalismo computacionalista concibe la cognición como procesamientos de información que implican la manipulación de representaciones, lo que siempre se efectúa siguiendo ciertas reglas. De acuerdo con ello, el cuerpo es simplemente un dispositivo de entrada/salida, que permite al sistema computacional, *i.e.*, al cerebro, adquirir nuevos datos de entrada por medio de los sentidos y, en consecuencia, éste seguirá sus órdenes para actuar y navegar en el entorno.

En este contexto, las ciencias cognitivas tradicionales priman la explicación acerca de que el cerebro es la única base para la realización de los procesos cognitivos, lo que ha implicado otorgar un papel trivial al cuerpo en los procesos de la cognición.

Pero ¿qué son las representaciones? en breve, en la teoría computacional de la mente, la *representación* es entendida frecuentemente como un objeto mental intencional (dirigido a algo en el mundo) con contenido y con propiedades semánticas, por tanto, con condiciones de corrección: nos dice cómo es el mundo, si nuestros estados mentales (p.ej. las creencias) en correspondencia son adecuados o inadecuados, etc., por lo que la representación es normativa⁴.

Entonces, dado lo anterior, el computacionalismo sostiene compromisos ontológicos y con diversas entidades teóricas que son muy claros: la cognición implica procesos de información y algorítmicos que operan sobre representaciones mentales. Dicha concepción de la cognición tiene consecuencias importantes al momento de explicar ciertos fenómenos mentales. Una consecuencia principal es que la cognición así definida resulta en una separación modal de la percepción y la acción, esto debido a que, según su explicación, el sujeto cognitivo se enfrenta con un mundo (pre-dado) que percibe a manera de *inputs*, puesto que la información proveniente de éste es la entrada para algún tipo de procesamiento cognitivo que opera de forma algorítmica, lo que produce la actualización de

⁴ La condición mínima para que un objeto mental cuente como representación es que sea una estructura portadora de información.

una serie de modelos y representaciones mentales que le permiten al sujeto arribar a la toma de decisiones y acciones (*outputs*). Así, prevalece una clara separación modal de la percepción y la acción en la ciencia cognitiva tradicional, tema que retomo más adelante⁵.

Por último, es conveniente hacer una aclaración de corte terminológico respecto a los términos “funcionalismo”, “computacionalismo” y “cognitivismo”, puesto que suelen ser utilizados como equivalentes. Dado que la versión funcionalista predominante en la ciencia cognitiva es el computacionalismo, es común que los términos “funcionalismo” y “computacionalismo” se utilicen en sinonimia tanto en la filosofía de la mente como en la ciencia cognitiva. No obstante, el término “cognitivismo” se emplea más en el campo de la psicología. Por lo que, en adelante, omito este último y utilizo en sinonimia los términos computacionalismo y funcionalismo.

Para mis propósitos, es suficiente con esta presentación panorámica de los supuestos cognitivos que se encuentran detrás del computacionalismo, puesto que son los que atraviesan las discusiones que se verán posteriormente.

Ahora, corresponde ver los supuestos existentes detrás del enactivismo.

1.3.2. Enactivismo

En la década de los 80 surge un diferente paradigma explicativo desde un programa nuevo de investigación en la ciencia cognitiva, específicamente, la cognición corporizada (Maturana y Varela, 1980; Varela, Thompson y Rosch, 1991; Steels y Brooks 1995; Clark 1997, 2008; Shapiro, 2011). En oposición fuerte a la ciencia cognitiva tradicional, la cognición corporizada no recurre a la representación mental, ni concibe que el cerebro sea la única base en la que se lleva a cabo la cognición, en su lugar, resalta la relevancia del cuerpo en relación con el ambiente en su explicación de los procesos cognitivos. En breve, la cognición es corporizada cuando depende en gran medida de las características del cuerpo físico de un agente, es decir, cuando los aspectos del cuerpo del agente, más allá

⁵ Cabe mencionar que, en enfoques funcionalistas posteriores a la propuesta computacionalista, surgió el “funcionalismo conexionista”, o bien, “el conexionismo” (en la década de los 80), cuyos fundamentos comparten algunos principios con el computacionalismo clásico, pero se separa de otros, por ejemplo, el conexionismo ya no habla estrictamente de un mecanismo regido por un sistema de reglas para explicar la cognición, en su lugar, la explica como un proceso que emerge de redes neuronales artificiales o de nodos. Sin embargo, sigue concibiendo la cognición como procesamientos de información que utilizan representaciones sub-simbólicas para su ejecución (esta mención es relevante para fines explicativos que se verán en el capítulo 4).

del cerebro, desempeñan un importante rol causal o constitutivo en el procesamiento cognitivo (Wilson y Foglia, 2017).

Ahora bien, en el marco de la cognición corporizada surge el programa del enactivismo. El enactivismo es un término general que se usa para describir varios enfoques relacionados dentro de la ciencia cognitiva y de la filosofía de la mente que, por lo general, enfatizan la naturaleza corporizada, dinámica y ambientalmente situada de la cognición. El punto de partida del enactivismo es explicar la cognición como acoplamientos sensoriomotores entre el agente y el entorno, en donde el agente *hace-sentido* del entorno, i.e. la cognición es entendida como interacciones agente-ambiente.

El enactivismo se considera una teoría naturalista, científica y no-reductiva de la mente, la cual surgió de un conjunto de ideas biológicas y fenomenológicas. El enactivismo ha presentado una poderosa crítica a la ciencia cognitiva tradicional respecto a sus nociones de cognición, autonomía y agencia por su enfoque individualista y racionalista de las mismas. Así, se ha sugerido una revisión de estas nociones en términos de una metodología *relacional* con énfasis en la naturaleza corporizada y situada de la cognición y de la agencia (Urban, 2014, 2015).

El enactivismo es un enfoque naturalista no-reductivo que propone una continuidad profunda entre los procesos de la vida y de la cognición (Varela *et. al.* 1991; Thompson, 2005, 2007; De Jaegher y Di Paolo, 2007; y Froese y Di Paolo, 2011). Es un programa científico que explora varias fases a lo largo de este continuo vida-mente, basado en los conceptos de autonomía, hacer-sentido (*sense-making*), corporización, emergencia, experiencia y hacer-sentido participativo (*participatory sense-making*). Estos conceptos forman un núcleo teórico que habilita las concepciones de la cognición enactiva.

Así, el enactivismo tiene como objetivo desarrollar un nuevo discurso que pueda integrar un conjunto diverso de criterios multidisciplinarios respecto a la interdependencia no lineal de los fenómenos a estudiar, puesto que, desde este enfoque, resulta imposible estudiar cualquier fenómeno, o un dominio de fenómenos, en completo aislamiento (Froese y Di Paolo, 2011). Con ello, ha desafiado un rasgo generalizado de la ciencia cognitiva tradicional, a saber, su individualismo metodológico; el supuesto de que el agente cognitivo individual es la unidad de análisis correcta para comprender la mente, así como también, su enfoque exclusivo en la agencia individual cuando se busca dar sentido al fenómeno social, lo que parece ser fundamentalmente erróneo desde la perspectiva enactiva. Tal

enfoque individualista de la mente y de la agencia ha sido reemplazado, desde el enactivismo, por un enfoque progresivamente creciente en la naturaleza interactiva y social de la experiencia y de la agencia (Torrance y Froese, 2011), i.e., por una metodología *relacional*.

De gran relevancia es que los teóricos enactivos enfatizan que los sistemas cognitivos, incluso en el nivel de formas de vida simples, “participan en la generación de significado a través de sus cuerpos y acciones que a menudo se involucran en interacciones transformacionales y no meramente informativas” (Di Paolo, *et al.*, 2010). Por lo que, un agente y el mundo significativo en el que éste actúa, deben considerarse como co-constitutivos o co-habilitadores (Torrance y Froese, 2011).

Entonces, en el enactivismo se trata de hablar de *acción* más que de *representación* y de un continuo entre la mente y la vida, más que de contenidos semánticos para habilitar la cognición.

En particular, considero virtuoso el énfasis en las dimensiones de la *interacción* y en el modelo *relacional* del que parte el enactivismo. Al respecto, argumento que las dimensiones de interacción y de mutualidad postuladas por el enactivismo tienen que ganar un lugar en el estudio de la cognición moral. Por ello, se verá más adelante que presento un análisis gradual respecto a los procesamientos cognitivos involucrados en el fenómeno a estudiar, esto con el objetivo de definir el alcance e implicación de cada uno de ellos.

Ahora, corresponde ver los principios de los que parte el enactivismo.

- **Principios del Enactivismo**

En el enactivismo la cognición es principalmente definida como acoplamientos sensoriomotores entre el agente y el ambiente, o sea, como interacciones agente-ambiente. En tales interacciones el ambiente se vuelve irreductiblemente significativo para el agente debido a que percibe sus características como oportunidades de acción, interacción y respuesta, constituyendo así una perspectiva y un punto de vista en el que los encuentros son significativos para el agente; el ambiente amenaza o permite la viabilidad de mantener

su integridad. Estas interacciones son la definición mínima del *hacer-sentido* (*sense-making*)⁶.

De esta manera, el enactivismo entiende la cognición como procesos de *hacer-sentido*, como la constante búsqueda o creación de sentido que caracteriza al agente en interacción con su entorno. Bajo esta definición, una propiedad crucial de los sistemas cognitivos es la *autonomía*, la cual refiere a un sistema constituido por varios procesos que activamente generan y sostienen una identidad bajo condiciones *precarias* (Di Paolo, 2005), teniendo así la capacidad para auto-mantenerse y también para auto-individuarse del ambiente, es decir, se define como una unidad que se auto-constituye y se distingue de su entorno. Ahora, que un sistema sea *precario* significa que se encuentra en constante riesgo de perecer debido a que está en desequilibrio, por lo que el sistema necesita estar en constante interacción con su entorno para intercambiar el flujo de materia y energía que le permita mantener su integridad (identidad) o, dicho de otro modo, el sistema es autónomo en tanto que produce y sostiene su integridad, lo que logra en interacción con su entorno. Entonces, el concepto de autonomía resalta un aspecto *individual*, en tanto que el agente se auto-constituye y se auto-mantiene, pero también un aspecto *relacional*, en tanto que logra preservar su integridad por la continua interacción con su entorno.

Ahora bien, la capacidad de generar y auto-mantener su identidad/integridad implica poseer la propiedad de *clausura operacional*, la cual especifica que entre las condiciones que habilitan la realización de cualquier proceso constituyente del sistema siempre se encuentra uno o más procesos que también pertenecen al sistema, es decir, no hay procesos que no estén condicionados por otros procesos en la red que conforma al sistema (Di Paolo, 2005). Por supuesto, esto no significa que procesos externos no influyan en procesos constituyentes del sistema, pero ellos no son parte de la red de relaciones de condicionamiento que existe entre los procesos constituyentes del sistema. De lo anterior se sigue que, sin la autonomía otorgada por la clausura operacional, el sistema sería incapaz de definir su propia identidad como individuo.

Otra característica clave del enactivismo es la propiedad de *adaptividad* del sistema autónomo, la cual refiere a la capacidad del sistema para *monitorear y regular activamente* sus interacciones con el entorno respecto a sus propias restricciones de viabilidad, o sea,

⁶ Se verá que este carácter exploratorio propio del hacer-sentido, también está presente en la cognición social enactiva, atribuyendo a los interactuantes el papel de *participantes* activos y no de meros *observadores*.

para favorecer su continuidad. En otras palabras, no se concibe como un sistema que pasivamente reciba información de su entorno, sino que el organismo *activamente* regula su relación con el medio, acción que se le puede llamar efectivamente *comportamiento*, debido a que tal regulación es efectuada por el organismo.

Así, sólo los sistemas *autónomos adaptivos* son capaces de la búsqueda y creación de sentido, *i.e.* de *hacer-sentido*, y con ello, de relacionarse de manera significativa (con valencia positiva o negativa) con su ambiente. De manera importante, el sistema evalúa su ambiente en función de la *normatividad* que surge de su propia auto-constitución en términos de una valencia dual: "atracción o repulsión, aproximación o escape" (Weber y Varela, 2002, p. 117). En otras palabras, a partir de la autonomía del sistema (mediante la cual el organismo se auto-individa, es precario y se auto-mantiene) se instaura una normatividad básica para el organismo que le permite mantener su identidad, o bien, para mantenerse vivo. Tal normatividad es intrínseca al organismo debido a que emerge de su propia organización interna y no de parámetros externos. A este nivel de normatividad se le llama *normatividad biológica*, la cual es individual o auto-dirigida y está basada en las necesidades metabólicas propias del sistema.

Por último, vale la pena ver la definición operativa de la "cognición" en la literatura enactiva:

"La cognición es el acoplamiento sensoriomotor regulado entre un agente cognitivo y su entorno, donde la regulación se dirige a aspectos del propio acoplamiento para que constituya una organización autónoma emergente en los dominios de la dinámica interna y relacional, sin destruir en el proceso la agencia de ese agente (aunque el alcance de este último puede aumentarse o reducirse)."

(Froese y Di Paolo, 2011, p. 18)

Esta breve presentación de ambos marcos conceptuales permite ver el panorama tan contrastante que prevalece entre ellos. En esta línea, a continuación, desarrollo las aportaciones que ofrece cada uno al estudio de la cognición social, para después, concentrarme en sus aportes hechos a la cognición moral humana.

Previo a ello, me interesa resaltar las consecuencias más relevantes de este apartado.

1.3.3. Consecuencias:

1. Para los computacionalistas hay una clara separación modal entre la percepción y la acción, en tanto que el sujeto percibe (*input*), procesa tal información percibida y

genera un resultado (*output*) en forma de toma de decisiones que después se convierten en acciones, mientras que, para los corporizados no hay tal separación, en su lugar, la cognición se efectúa a manera de bucles continuos de percepción \leftrightarrow acción, lo que propicia más posibilidades de actuación y significación para el agente.

2. Los computacionalistas conciben la cognición como procesamientos de información regidos por reglas y los corporizados como procesos de hacer-sentido habilitados por las interacciones agente-entorno.
3. Los computacionalistas parten de un individualismo metodológico (en tanto que su principal unidad de análisis es el sujeto individual y los procesos intracraneales que en éste ocurren), mientras que el enactivismo parte de una metodología relacional que suele defender una epistemología interaccionista.
4. El enactivismo habla de *acción* y de un continuo entre la mente y la vida, mientras que el computacionalismo trata con *representaciones* y contenidos semánticos para explicar la cognición.

Así, quedan de relieve las divergencias teóricas que predominan entre los dos marcos conceptuales. Ahora, corresponde ver las aportaciones que cada uno ofrece a la cognición social.

1.4. Cognición social

En general, el término “cognición social” refiere a las capacidades que habilitan nuestras diversas maneras de entender y de interactuar con otros, así como también, de aquellas competencias que nos permiten navegar exitosamente en el mundo social.

El tema de la cognición social resulta esencial para mi propuesta, por lo que es relevante presentar el panorama conceptual en el que se sitúa. Para ello, explico de manera general las dos concepciones antagónicas de las competencias socio-cognitivas: desde los principios del funcionalismo-computacionalista, por un lado, y desde los principios corporizados-enactivos, por el otro. En breve, el debate en la cognición social es entre atribucionistas vs corporizados.

Del lado de los atribucionistas (computacionalistas), me enfoco en sus dos principales teorías: teoría-teoría (Leslie, 1987, 1994; Baron-Cohen, 1995; Gopnik, 1990, 1996; Wellman, 2002) y teoría simulacionista (Gordon, 1996; Goldman, 2006; Heal, 1996). Y, respecto a las ciencias cognitivas corporizadas, resalto su aspecto interaccionista y explico

sus tres enfoques principales: 1) la teoría de la interacción (Gallagher, 2001, 2004); 2) el enfoque de la segunda persona (Hutto, 2004; Reddy, 2008; Ratcliffe, 2007), y; 3) el enfoque enactivo (De Jaegher, 2009; De Jaegher y Di Paolo, 2007). Comencemos con los atribucionistas.

1.4.1. Atribucionistas

En la década de los 80 surgieron las dos teorías filosóficas pioneras (teoría-teoría y teoría simulacionista) que intentaron explicar la competencia socio-cognitiva humana, centrándose en clarificar cómo logramos interpretar las conductas y las acciones de otro individuo, teorías que hoy se les conoce como "teorías de la atribución mental" o "teorías de lectura-de-mentes (*mindreading*)".

El postulado principal de los atribucionistas es que los seres humanos adscribimos estados mentales a nuestros congéneres con la finalidad de describir, comprender, explicar y predecir sus conductas. Sus dos modelos teóricos (teoría-teoría y teoría simulacionista) son el punto de partida de numerosas y considerables investigaciones, tanto desde el punto de vista filosófico como psicológico y neurocientífico, lo que ha constituido el robusto basamento teórico-empírico que apuntala a cada uno de ellos. Sin embargo, la cuestión acerca de cómo opera nuestra capacidad de adscripción mental, o bien, cómo logramos atribuir estados mentales a los demás en orden a predecir y explicar sus conductas, es materia de debate entre la teoría-teoría (TT en adelante) y la teoría de la simulación (TS en adelante)⁷.

Enfáticamente, el eje de debate entre la TT y la TS es si la adscripción mental depende de una capacidad de teorización implícita, o bien, del despliegue de un conjunto de habilidades simulativas que poseemos los seres humanos, pues ambas teorías parten del supuesto de que todas las personas, sin entrenamiento o enseñanza alguna, logramos un rango moderado de precisión atribucional en nuestras prácticas predictivas-explicativas cotidianas, por tanto, la pregunta se centra en ¿cómo lo logramos? ¿es mediante una capacidad simulativa, o bien, es a través del uso de una teoría tácita?

De esta manera, el debate computacionalista se ha articulado en torno a explicar la naturaleza de la cognición social humana en términos de atribución mental, ya sea teórica

⁷ Cabe mencionar que actualmente existen enfoques híbridos que consideran que los dos procesamientos trabajan en conjunto para lograr una atribución mental adecuada (Nichols y Stich, 2003; Heal, 1994; Goldman, 2006).

o simulativa. A continuación, presento los detalles de las explicaciones de la TT y posteriormente los de la TS.

- Teoría-Teoría

Los teóricos de la TT sostienen que las personas legas construyen o son dotadas con una teoría psicológica tácita del sentido común o "psicología folk" (PF en adelante) que guía sus asignaciones de estados mentales, es decir, desde este enfoque la PF es considerada como una teoría que funciona para atribuir, interpretar, explicar y predecir las conductas de los demás. Los detalles de cómo opera la PF, si es innata o adquirida, es materia de debate entre nativistas y empiristas.

Los teóricos nativistas, apegados en su mayoría a los principios evolucionistas de la escuela de Santa Bárbara, plantean que en el cerebro existen módulos especializados en computar la información de cada una de las distintas capacidades cognitivas, tales módulos son innatos y se desarrollan en función de la maduración biológica, de esta manera, existe un módulo especializado para la teoría de la mente (ToMM por sus siglas en inglés *Theory of Mind Module*). La idea es que tenemos una capacidad modular convertida en una facultad de calcular o *computar* mentes, es decir, un mecanismo de especial atención que emplea conceptos innatos tales como creencias y deseos, conceptos innatamente específicos que predisponen al desarrollo normal de los infantes a prestar atención selectiva a los estados mentales, lo que explicaría cómo el infante aprende acerca de estados mentales y cómo atribuirlos (Fodor, 1983, 1992; Leslie, 1987, 1994; Baron-Cohen, 1995). Este enfoque lo retomo en el capítulo 2.

Por su parte, los empiristas defienden la "teoría del niño científico" (Gopnik 1990, 1996; Carey y Spelke, 1996; Meltzoff 1999; Wellman y Gelman 1997; Wellman 2002) y sostienen un innatismo moderado de las capacidades o tendencias del sujeto para realizar cierto tipo de actividades psicológicas, las cuales son desarrolladas mediante aprendizaje o "teorización".

Los teóricos empiristas sostienen que la comprensión psicológica de los infantes se estructura y funciona de la misma manera que las teorías científicas, es decir, los infantes explican la conducta propia y ajena usando teorías formadas por constructos teóricos representacionales sobre estados mentales. La propuesta es que hay una comprensión temprana de deseos y emociones, pero no de creencias, dado que los datos y la información

se acumulan para conducir a un cambio de teoría, donde posteriormente se incorpora la comprensión de creencias dentro de las explicaciones psicológicas.

Así, en analogía con las teorías científicas, los niños construyen teorías de la mente cada vez más complejas y elaboradas, las cuales experimentan cambios por ser revisadas y probadas para después rechazarlas, reemplazarlas, modificarlas o ampliarlas. Lo que motiva el planteamiento es que las adscripciones de estados mentales proceden mediante inferencias que son guiadas por principios causales de la PF y tales principios surgen de la construcción y revisión de la teoría como sucede en la construcción y revisión de las teorías científicas.

Ahora bien, pese a estas discrepancias entre las versiones TT, tanto nativistas como empiristas comparten los siguientes requerimientos teóricos funcionalistas que habilitan su modelo explicativo: 1) la atribución de estados mentales a otros se efectúa mediante razonamiento teórico, cuyas premisas son las características observables de la conducta y del entorno, más, la aplicación de reglas o principios causales de la PF; 2) los conceptos de estados mentales son entendidos en términos de relaciones causales teóricamente específicas entre conducta, medio ambiente y otros estados mentales, es decir, son entendidos en términos de sus roles causales o funcionales; 3) el desarrollo de las habilidades de atribución mental se da en función de los cambios ocurridos en la teoría psicológica folk del atribuidor (ya sea por maduración biológica, como lo defienden los innatistas, o bien, por aprendizaje, como lo sostienen los empiristas).

En concreto, la TT arguye que para interpretar, explicar y predecir las conductas de los demás, recurrimos a una teoría tácita de la PF que los seres humanos poseemos representada en nuestra mente-cerebro (Ravenscroft, 2004) -ya sea por dotación biológica o por aprendizaje-, la cual se constituye de un amplio repertorio conceptual (donde las actitudes proposicionales paradigmáticas son los deseos y las creencias), y de un conjunto de principios o generalizaciones legaliformes del sentido común que relacionan causalmente a los estados mentales con otros estados mentales, con estímulos del medio ambiente y con la conducta; asimismo, la TT postula que los estados mentales preeminentemente poseen un carácter causal en virtud de su contenido, por lo que deben ser conceptualizados en términos de esos roles causales o funcionales y, por último, que su adquisición y/o desarrollo depende de que el infante los comprenda como constructos representacionales que relacionan el mundo mental con la realidad externa.

- Teoría Simulacionista

En contraste, la TS sugiere que el razonamiento ordinario acerca de personas y mentes no se da por constructos teóricos, sino que los estados mentales primeramente se experimentan en uno mismo y después éstos se atribuyen a las otras personas mediante un proceso denominado *simulación*. La idea es que se hace uso de una estrategia que consiste en *ponerse imaginativamente en el lugar del otro* para comprender sus acciones y lograr así la atribución mental; el atribuidor se imagina a sí mismo en las mismas circunstancias del otro, simulando los mismos deseos, creencias o emociones que se introducen en su propio mecanismo de toma de decisiones y la conclusión resultante se la atribuye al sujeto de adscripción, con esto explica y predice su comportamiento, llevándolo a cabo sin imputar al agente simulador un cuerpo de conocimientos generales implícitos sobre leyes o generalizaciones psicológicas. Lo fundamental de la TS es que los atribuidores utilizan los recursos de sus propias mentes para simular o imitar la mente ajena y de este modo explicar o predecir las conductas del sujeto de la adscripción.

En particular, la TS en su mayoría descansa en los principios de similitud, empatía, racionalidad, procesos *off-line*⁸ y proyección (a continuación, se verán los detalles de cada uno de ellos), de lo cual, emergen sus principales requerimientos (Gordon 1986; Heal 1986; Goldman 1989; Harris, 1991, aunque entre sus tesis existen diferencias significativas).

Los requerimientos teóricos que habilitan a la TS son: 1) los sistemas (internos y/o externos) del atribuidor y del sujeto de adscripción deben ser lo bastante similares para que el primero pueda ser modelo del segundo. Este requerimiento proviene del principio de *similitud*, el cual refiere a que en la práctica simulacionista se asume que el sistema del simulador y el sistema del simulado mantienen una relación de *similitud* para que el primero pueda ser modelo del segundo y así lograr, por analogía, resultados fiables y exitosos. Por tanto, los simulacionistas sostienen que para adscribir estados mentales a otros "simplemente experimentamos y reportamos nuestras propias experiencias mentales para después poder atribuir tales experiencias a los demás" -en Goswami, 2004, p. 184-); 2) el atribuidor toma

⁸ Procesos *off-line* implican representaciones internas, las cuales no están vinculadas con las características actuales del cuerpo del agente o de su entorno, por lo que tales representaciones son capaces de soportar situaciones contrafácticas. Pertinente a la adscripción mental simulacionista, el planteamiento es que el sistema simulador recoge las entradas hipotéticas simuladas (*inputs*) desconectando paralelamente los sistemas de entrada y central del sistema de salida, desligando así los mecanismos de control de la acción, de manera que los datos resultantes (*outputs*) del proceso simulacionista no afecten el sistema del simulador y éste pueda libremente atribuírselos al sujeto de adscripción.

la perspectiva del otro mediante su capacidad empática; éste refiere al principio de *empatía*, que es el cimiento básico del ejercicio simulacionista, dado que es la habilidad subyacente a uno de los principales requerimientos de la TS, a saber, “la toma de perspectiva de otro”, pues los humanos manifestamos poseer una capacidad empática, capacidad que nos permite situarnos imaginativamente en el lugar del otro para sentir o experimentar sus sensaciones o experiencias como si fueran nuestras. De esta manera, en términos simulacionistas, el atribuidor tiene la capacidad de adoptar una perspectiva espacio-temporal o situacional diferente a la suya; 3) en la práctica cotidiana de adscripción mental asumimos que el otro es un agente racional, es decir, le suponemos racionalidad al sujeto de adscripción. En específico, este requerimiento proviene del principio de *racionalidad*, que refiere al hecho de que en la práctica predictiva-explicativa de atribución mental se presupone que las personas actúan racionalmente conforme a una competencia cognitiva compartida e inmutable con carácter ubicuo (esta afirmación es altamente debatible, por lo que más adelante dedico unas palabras al respecto); 4) en la adscripción mental a terceros es requerimiento el uso, en modo *off-line*, de los sistemas que controlan la conducta del atribuidor; y, 5) la proyección es parte de lo que consiste en simular correctamente los estados mentales del otro. La *proyección* ocurre cuando le adscribimos el estado mental simulado a otra persona, ya que una característica fundamental de la TS es que el atribuidor utiliza su propia mente para adscribir estados mentales a otros, esto es, toma sus propios estados simulados y se los atribuye al sujeto de adscripción como estados genuinos.

En este punto son pertinentes unas palabras acerca del principio de *racionalidad*, pues alude a una cuestión altamente controversial. Al respecto, los teóricos de la atribución mental que recurren a esta noción de *racionalidad* se abocan básicamente a lo que Richard Grandy (1973) denominó el *principio de humanidad*. En breve, tal principio apela a la similitud de los patrones de relaciones entre los deseos, las creencias y el mundo que hay entre los sujetos, es decir, nos suponemos similitud entre los sujetos para efectuar las atribuciones mentales, al menos en determinadas circunstancias. A estos fines menciona Grandy:

“Entonces, es de fundamental importancia efectuar las interrelaciones entre esas actitudes del modo más similar posible a las nuestras. Si la traducción nos dice que las creencias y los deseos de la otra persona están conectados de un modo demasiado extraño como para que nosotros le demos sentido, entonces, la traducción no sirve para nuestros fines. Así, tenemos como un requisito pragmático de la traducción, la condición de que los patrones imputados de relaciones entre los

deseos, las creencias y el mundo sean lo más similares posibles a los nuestros. Llamaré a este principio el principio de humanidad” (Grandy, 1973, p. 443).

Estrictamente hablando, algunos autores como Edward Stein consideran adversa y problemática la cuestión de la racionalidad humana. *Grosso modo*, Stein (1996), basado en evidencia empírica, plantea que los seres humanos no siempre razonamos correctamente (de hecho, afirma que la mayoría de las veces razonamos incorrectamente), puesto que presentamos errores significativos en diversos campos del razonamiento, tales como razonamiento lógico y probabilístico entre otros, por lo que sugiere que se debe cambiar la noción de *racionalidad* a la que comúnmente se apela en filosofía y en ciencia cognitiva. En esta línea, en el campo de la psicología se proponen las teorías duales del razonamiento, las cuales plantean dos sistemas cognitivos de razonamiento: uno que responde rápidamente y casi de manera automática, es decir, son formas de razonar no formales; y el otro, correspondiente al razonamiento lento y consciente. Así pues, existe un álgido debate en torno a la noción de la racionalidad humana, por lo que debe resaltarse que los teóricos de la atribución mental se ciñen a la concepción de *racionalidad* mencionada, es decir, al *principio de humanidad* de Grandy.

Ahora bien, hay dos versiones prominentes en la TS: la “introspeccionista” y la “no-introspeccionista”. La primera propone que la atribución de estados mentales a otros depende completamente de nuestro acceso introspectivo a nuestros propios estados mentales, aunado a procesos inferenciales, analógicos y de simulación (Goldman, 1993, 2006); y la segunda enfatiza que la atribución mental se efectúa en función de que todos vemos uno y el mismo mundo, por tanto, la adscripción depende de mirar al mundo y no del aspecto introspectivo (Gordon, 1986, 1996, 2000).

La versión no-introspeccionista de Robert Gordon, quien es un defensor del simulacionismo radical, sostiene que nuestra habilidad de adscripción de estados mentales requiere de un “cambio egocéntrico” o “recentramiento” del mapa egocéntrico del atribuidor al sujeto de adscripción. Tal “cambio egocéntrico” refiere a una “transformación” de mí mismo (atribuidor) en ti (objetivo), es decir, imaginativamente, el pronombre “yo” se vuelve “tú”, donde “yo” no pienso, no creo y no deseo como “yo”, sino que, ahora pienso, creo y deseo como “tú”, por tanto, el autor no apela a inferencias analógicas de “mí” a “ti”, sino que, en su lugar, habla de una simulación entendida como “transformación”.

Otro elemento fundamental en la teoría de Gordon es que presupone una similitud no-interna entre atribuidores y sujetos de adscripción, es decir, una similitud externa entre ellos,

la cual es una similitud de un mundo compartido y mutuamente conocido, donde lo relevante para la adscripción de estados mentales es el acceso a ese mismo mundo de hechos objetivos que dirige a los sujetos a la acción.

Así, la similitud externa es el fundamento habilitador de su hipótesis “rutinas ascendentes”, cuya idea básica es que las tareas de atribución mental son rutinarias y no requieren necesariamente de razonamiento para efectuarlas, pues dichas adscripciones se hacen en virtud de los hechos que suceden en el mundo. El autor las nombra “ascendentes”, debido a que los estados mentales de las personas provienen del estado fáctico del mundo; idea que hace alusión a un proceso ascendente: desde el estado del mundo, hasta el estado mental de la persona⁹.

Por su parte, la versión introspeccionista de Alvin Goldman se compromete con tres supuestos teóricos para efectuar la atribución mental: 1) inferencia analógica de uno mismo a otro; 2) adscripciones de estados mentales basadas introspectivamente en uno mismo, y; 3) la posesión previa de los conceptos de los estados mentales a adscribir. Tomando esto como base y valiéndose de resultados neurocientíficos y psicológicos, Goldman (2006) postula un sistema bi-nivel simulacionista para dar cuenta de la atribución mental: a) lectura-de-mentes de bajo-nivel y b) lectura-de-mentes de alto-nivel.

La lectura-de-mentes de bajo-nivel se encarga de la lectura de estados cognitivos básicos como las emociones básicas (miedo, repulsión e ira), la cual procede de manera automática y no necesariamente consciente mediante imitación mental como simulación interpersonal (cuando los procesos mentales X y X' ocurren en diferentes mentes), en la que el sistema de neuronas espejo (NE en adelante) juega un papel crucial, pues, según el autor, es el que nos permite entrar en los mismos estados mentales que observamos en otra persona, logrando así la atribución mental.

La lectura-de-mentes de alto-nivel, por su parte, se encarga de la lectura de estados cognitivos de nivel superior como las actitudes proposicionales (deseos, creencias e intenciones) y procede con algún grado de accesibilidad a la consciencia debido a que algunos de los componentes de su proceso están sujetos a control voluntario, efectuándose

⁹ Una restricción que se debe tener en cuenta, es que, no por tomar la lógica de *transformación* se le otorga al atribuidor un papel de *participante activo* y, además, este enfoque continúa con la suposición teórica de que la atribución mental es el método principal para entender e interactuar exitosamente con los demás. Esto lo explico más adelante, en donde abordo los contrastes entre los atribucionistas y los corporizados.

mediante pretensión mental o *imaginación recreativa*¹⁰, dado que ésta es postulada como el recurso para lograr un alto rango de precisión atribucional.

De acuerdo con lo anterior, el sistema bi-nivel goldmaniano funciona de la siguiente manera:

Lectura-de-mentes simulacionista de bajo-nivel: en función de NE, cuando un sujeto A observa en el rostro de un sujeto B la expresión de una emoción X, automáticamente se desencadenan, en el sistema del sujeto A, los mismos sustratos neuronales implicados en la experimentación de la emoción X, tal y como si él mismo se encontrara experimentando la emoción observada, entonces, el sistema emocional del sujeto A "imita" o "resuena" con el del sujeto B, generando así, la atribución de la emoción X al sujeto B.

Lectura-de-mentes simulacionista de alto-nivel: el atribuidor observa al otro y, en función de la *imaginación recreativa*, simula el estado p observado, el cual es suficientemente semejante a su contraparte de estado genuino no-simulado p' , con lo que logra una atribución mental de alto-nivel precisa o exitosa.

Ahora bien, pese a que TT y TS están en desacuerdo acerca de los procesos y mecanismos que efectúan la atribución mental, su relación teórica es que ambas asumen que el modo principal y fundamental que utilizamos para lograr comprender e interactuar con nuestros congéneres es la adscripción de estados mentales, mediante lo cual explicamos y predecimos sus conductas, siendo éste el supuesto asumido en los últimos 40 años que ha dominado la literatura de la cognición social; supuesto que me interesa cuestionar desde una perspectiva corporizada.

La importancia de tener en cuenta el debate computacionalista entre la TT y la TS, es que en la actualidad hay propuestas híbridas constituidas por ambos métodos de adscripción mental (Nichols y Stich, 2003; Heal, 1996; Goldman, 2006), es decir, ambos métodos se traslapan en un marco integral en el que las inferencias teóricas de la TT y las simulaciones de la TS (tal como se describieron) trabajan de manera independiente o co-participando en tareas de atribución mental, por lo que es crucial comprender cómo funciona cada una.

En este contexto, son las propuestas híbridas las que permiten ampliar el marco de debate. En palabras de Hutto:

¹⁰ *Imaginación recreativa* en el sentido de que es un tipo de imaginación capaz de representar los estados mentales del otro y de recrear una situación dada (ver Goldman, 2006).

“Después de años de ordenar las consideraciones filosóficas y la evidencia de la psicología, quedan pocos puristas de TT/TS; la mayoría de los investigadores que trabajan en psicología cotidiana aceptan que la mejor esperanza de proporcionar una descripción adecuada de las habilidades relevantes se basará en los recursos de ambos campos”. (Hutto, 2009, p.223)

Por esta razón me interesa tener claros los debates conceptuales de la cognición social y los recursos que cada postura emplea. Hasta aquí se han visto las propuestas computacionalistas, ahora corresponde ver las propuestas corporizadas.

1.4.2. Cognición social corporizada

En este marco de debate surge, en el año 2001, un giro conceptual llamado “giro interactivo en la investigación de la cognición social” proveniente de las críticas efectuadas a las explicaciones atribucionistas. En general, las críticas se enfocan en acusar a las perspectivas atribucionistas de ignorar el papel de la *interacción social* en sus explicaciones y, en su lugar, enfocarse exclusivamente en el agente cognitivo individual y en los procesos que ocurren dentro de su cabeza para navegar en el mundo social.

Shaun Gallagher es el pionero del giro interactivo y como punto de partida propone la “teoría de la interacción” (2001). De acuerdo con su propuesta, las teorías atribucionistas son deficientes debido a que han ignorado el papel fundamental de los contextos interactivos en los que tiene lugar la cognición social, así como también, las intervenciones corporizadas que la habilitan. Estas críticas han tenido un fecundo desarrollo en la literatura corporizada de la cognición social, lo que ha resultado en diversos enfoques y explicaciones.

No obstante, el principal postulado que comparten estos enfoques es que la *interacción* es el recurso de la intersubjetividad, de ahí que se les ha etiquetado como “el enfoque *interaccionista*”. Por supuesto, también se les unifica por tomar como punto de partida las críticas que se le han hecho a las teorías atribucionistas que comprende el giro interactivo.

Ahora bien, sus principales enfoques explicativos son: 1) “la teoría de la interacción” (Gallagher, 2001, 2004); 2) “el enfoque de la segunda persona” (Hutto, 2004; Reddy, 2008; Ratcliffe, 2007), y; 3) “el enfoque enactivo” (De Jaegher, 2009; De Jaegher y Di Paolo, 2007). Cada uno de ellos sostiene compromisos filosóficos que deben ser explicados.

Comienzo con “la teoría de la interacción” de Gallagher, la cual, pese a que ha sido blanco de críticas, también se le ha considerado como el punto de partida de un campo fértil de

investigación en el área, esto por poner en primer plano el *contexto interactivo* en el estudio de la cognición social.

Previo a continuar, cabe mencionar que “el enfoque de la segunda persona” captura las bases que habilitan a cualquier explicación corporizada, pues existe un álgido debate respecto a las perspectivas epistemológicas de primera, de segunda y de tercera persona que deriva en implicaciones cruciales en las concepciones de la cognición social, por lo que en esta sección abordo dicho debate. No obstante, en el capítulo 3 se verá que esta perspectiva se legitima a través de un cuerpo robusto de evidencia empírica

- **Teoría de la interacción de Gallagher**

Gallagher sugiere que la cognición social no requiere, en primer lugar, ni de las capacidades inferenciales postuladas por la TT, ni de las capacidades simulativas postuladas por la TS, sino, más bien, requiere de capacidades más básicas y facilitadas por la *percepción directa* de los estados mentales del otro. Según Gallagher, nuestra vida social se negocia esencialmente a través de interacciones en segunda persona que se ven facilitadas por una comprensión perceptiva directa de los comportamientos corporales expresivos de los demás, lo que constituye un andamiaje de posibilidades de acción y respuesta (ver van Grunsven 2018, pp. 137-138).

En esta línea, Gallagher postula una explicación de la cognición social que no recurre a la atribución mental, sino que, en su lugar, recurre a prácticas corporizadas más básicas para explicar cómo logramos entender e interactuar en contextos sociales. Estas prácticas corporizadas están constituidas por lo que el autor llama *intersubjetividad primaria* e *intersubjetividad secundaria*, las cuales constituyen nuestra principal forma para lograr entender e interactuar en entornos sociales.

La *intersubjetividad primaria* es el entendimiento corporizado, pre-teórico, no-conceptual de otros agentes que subyace y fundamenta las habilidades cognitivas de alto-nivel (como las postuladas por la TT y la TS); "es el desarrollo de una capacidad innata o temprana para interactuar con otros manifestada en el nivel de experiencia perceptual -nosotros vemos o, más general, percibimos en los movimientos corporales, gestos faciales, dirección de la mirada, etc., de otras personas, lo que ellos pretenden y lo que ellos sienten" (Gallagher, 2004, p. 204).

La *intersubjetividad primaria* consiste en la sensibilidad informacional y la capacidad de reacción apropiada a características específicas del entorno. Estas habilidades están presentes en la primera infancia y permiten el entendimiento social y la interacción social. Se refiere, entre otras cosas, a la imitación neonata (Meltzoff y Moore, 1977), a la participación en proto-conversaciones en la infancia (Trevarthen, 1979), a la sintonización afectiva (Stern, 1985), al seguimiento de la mirada (Senju *et al.*, 2006) y a la percepción de movimientos intencionales o relacionados a objetivos a los 9 meses de edad (Meltzoff, 1995; Csibra, *et al.*, 1999), así como también, a la capacidad de completar una acción incompleta de un sujeto observado a los 18 meses de edad (Herrmann *et al.*, 2007; Meltzoff, 1995).

La *intersubjetividad secundaria*, por su parte, se desarrolla alrededor del primer año de edad y se efectúa por un movimiento de uno-a-uno; es una intersubjetividad inmediata a los contextos de atención compartida. En adición al seguimiento de la mirada, a la detección de la conducta intencional y a la percepción de las emociones de otros, con el desarrollo de la *intersubjetividad secundaria*, el infante desarrolla la capacidad para participar en comportamientos de atención compartida. "La característica definitoria de intersubjetividad secundaria es que un objeto o un evento puede convertirse en un foco entre las personas. Objetos y eventos pueden comunicar acerca de asuntos [...]. Las interacciones de los infantes con otras personas comienzan a tener referencias a las cosas que los rodean" (Gallagher, 2004, p. 207). Bajo la interpretación de la psicología del desarrollo, en esta etapa el infante aprende a seguir las miradas, a señalar y a comunicarse con otros acerca de objetos de atención compartida. Con ello, la capacidad del niño para el entendimiento social se desarrolla ampliamente.

En suma, la teoría de la interacción sostiene que, para desarrollar la capacidad para poseer creencias acerca de los estados mentales de otros, primero debe haber una comprensión de estas prácticas corporizadas más básicas, pues es mediante ellas como se va adquiriendo el significado de los encuentros cotidianos con los demás.

- **El enfoque de segunda-persona**

En el marco de la cognición social, hay un debate prevaeciente entre sus concepciones, el cual es respecto a lo que se conoce como "la perspectiva de primera persona", "la perspectiva de segunda persona" y "la perspectiva de tercera persona".

La perspectiva de primera persona sostiene que el entendimiento de los otros agentes se basa en las experiencias propias: primero *yo* experimento el estado mental relevante para después atribuírselo al *otro*. Esta perspectiva es el recuso común de los simulacionistas (Goldman, 2006; Heal, 1996).

Por su parte, la perspectiva de segunda persona refiere a los encuentros *cara-a-cara* que aborda la literatura de la cognición corporizada, la cual es favorecida por un entendimiento perceptual directo del otro (Reddy, 2008).

En contraste, la perspectiva de tercera persona es de *observador* y se relaciona con un punto de vista objetivo para lograr el entendimiento de otros. Esta perspectiva es el principal recurso de las teorías atribucionistas y ha sido acusada de ser “afectivamente desvinculada” en la literatura corporizada (Ratcliffe, 2007).

Pero, ¿cuál es la relación concreta entre estas perspectivas epistemológicas y las explicaciones de la cognición social?

En particular, la TT toma una perspectiva de tercera persona que refiere al hecho de que, si la *psicología folk* es una teoría, entonces las atribuciones que se den en virtud de ella parten de la perspectiva de tercera persona. Explico, la perspectiva de tercera persona es considerada como materialista, exterior y perteneciente al mundo objetivo, es decir, de acuerdo con ella, la TT utiliza el método de la *observación* externa de las conductas con la finalidad de revelar la estructura causal de los procesos mentales, esto por seguir los principios funcionalistas. En esta línea, la TT sostiene que hay una *simetría* funcional entre la atribución mental de primera persona y de tercera persona (*i.e.* la atribución mental a terceros no requiere de la auto-atribución, sino que la primera puede efectuarse sin la auto-atribución del estado mental en cuestión), afirmación que entra en debate con la TS, pues ésta rechaza tal *simetría* y defiende que la auto-atribución necesariamente antecede a la atribución mental a terceros.

En este contexto, la TS se basa en una perspectiva de tercera persona, pero también se beneficia de la perspectiva de primera persona; por sostener que la auto-atribución mental es previa a la atribución mental a terceros (requiere del reconocimiento del *yo* en el otro), con lo que defiende una *asimetría* funcional entre la perspectiva de primera y de tercera persona. Para ser clara, la rutina TS se basa en la perspectiva de *observador*, *i.e.* de 3ª persona, pero, a su vez, toma como recurso la perspectiva de 1ª persona debido a que

sostiene la idea de que “para poder atribuir un estado mental, primero debí experimentar yo ese preciso estado mental”. Esto puede quedar más claro si advertimos el proceso de la atribución mental de la TS: su punto de partida es observar al otro (perspectiva de 3ª persona), después recurre a la perspectiva de 1ª persona mediante su método de “me pongo imaginativamente en tu lugar” y, por último, el estado mental resultante se lo atribuye al otro (otra vez, desde la perspectiva de 3ª persona)¹¹.

Así, la TS, si bien recurre al recurso de la perspectiva de 1ª persona, debe quedar claro que se fundamenta en la perspectiva de 3ª persona (de observador) para lograr la atribución mental. Los debates tocantes a la *simetría/asimetría* entre la TT y la TS son extensos y con diversos matices¹², lo que sale del alcance del presente trabajo, sin embargo, para mis propósitos, es suficiente con tener claro que tanto la TT como la TS se fundamentan en la perspectiva epistemológica de 3ª persona, i.e. de *observador*, para explicar la cognición social.

En contraste, las explicaciones corporizadas parten de una perspectiva epistemológica de segunda persona y con ello otorgan un papel de *participante activo* y no de mero *observador* a los agentes que interactúan, lo que tiene implicaciones cruciales en el estudio de la cognición moral. En palabras de Di Paolo *et. al.*:

“Debido a que un enfoque enactivo otorga gran importancia a la autonomía de los individuos involucrados [en la interacción], este enfoque de la cognición social, si bien se centra en el proceso de interacción, paradójicamente también otorga a los agentes sociales una autonomía y un papel que no ha sido tematizado antes: el de participación en contraste con la mera observación” (Di Paolo *et. al.*, 2010, p. 72).

Así, pues, las implicaciones cruciales son que las explicaciones computacionalistas de la cognición social (TT y TS) parten del individuo para explicar la interacción desde una perspectiva de tercera persona (en la que los interactuantes son meros *observadores*), mientras que las explicaciones socio-cognitivas corporizadas parten de la perspectiva epistemológica de segunda persona, las cuales son facilitadas por la percepción directa y su recurso habilitador es la interacción misma.

¹¹ Cabe mencionar que tal combinación de perspectivas epistémicas ha beneficiado los enfoques híbridos de la cognición social computacionalista.

¹² Para una discusión detallada al respecto ver Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32, pp. 121–182.

De esta manera, desde un punto de vista enactivo, se puede hablar del surgimiento de la cognición social a través de la interacción (diádica) en segunda persona, de tal suerte que:

- 1) Las interacciones diádicas son parte integral del desarrollo de una amplia gama de habilidades que nos permiten identificar y responder de manera confiable a las intenciones, deseos y necesidades de los demás.
- 2) Las interacciones coordinadas representan significados sociales que no se pueden reducir a las intenciones de los agentes individuales.

(ver Grunsvén, 2018, p. 9)

Cabe subrayar que las afirmaciones de la perspectiva de 2ª persona se apoyan en un cuerpo robusto de evidencia empírica, la cual expongo en el capítulo 3.

Por el momento, corresponde destacar los supuestos de los que parte el enactivismo para dar cuenta de la cognición social.

- **Enfoque enactivo**

El enactivismo se basa en los supuestos de la perspectiva de 2ª persona y su punto de partida es la noción enactiva de hacer-sentido (*sense-making*), pero ahora enfocado al nivel de la interacción entre dos o más agentes autónomos, a lo que Di Paolo y De Jaeguer (2007) han llamado *hacer-sentido participativo* (PSM en adelante por sus siglas en inglés *participatory sense-making*).

Un movimiento conceptual clave del PSM es que los autores usan el principio de *autonomía*, pero aplicado al proceso *relacional* que toma lugar en el encuentro de dos o más agentes, es decir, la interacción es *autónoma* en sí misma. Esto significa que la interacción constituye un dominio cognitivo capaz de habilitar y constreñir la conducta individual, así como también, da lugar a una mirada de significados compartidos y normatividades co-constituidas por los agentes que participan en ella; movimiento conceptual que instala un genuino dominio *relacional* irreductible a la cognición individual. De esta manera, un PSM genuino exige el reajuste frecuente del hacer-sentido individual de los interactuantes a partir de acoplamientos dinámicos entre ellos, en cuyo proceso no debe ser destruida la autonomía de los participantes. Los autores lo describen de la siguiente manera:

"hacer-sentido participativo... es la coordinación de la actividad intencional en la interacción, mediante la cual los procesos individuales de hacer-sentido se ven afectados y se pueden generar nuevos dominios de hacer-sentido social que no estaban disponibles para cada individuo por sí solo. [...] hacer-sentido participativo no se puede reducir a acciones o intenciones individuales, sino que instala un

dominio relacional con sus propias propiedades que limita y modula el comportamiento individual " (Di Paolo y De Jaeguer, 2007, p. 497).

Crucialmente, en el caso de los seres humanos "la sintonía perceptiva y motora interactiva con el mundo [...] es principalmente con un entorno de [...] las acciones intencionales de otros" (Thompson 2007, p. 80). Lo que a su vez significa que: "los agentes humanos, ante todo, enactúan, sostienen y construyen sus precarias vidas autónomas a través de una sintonía perceptiva y motora con las acciones intencionales de los demás [...]" (van Grunsven 2018, p. 137).

Así, pues, la naturaleza precaria de la autonomía humana requiere de la continua interacción con su entorno y con los demás, lo que denota su condición de *dependencia recíproca*, tema que retomo en el capítulo 3.

Por el momento, con estos principios es posible comenzar a ver cómo en la literatura de la cognición social corporizada se le otorga un papel de *participante* involucrado (y no de mero *observador*) a los agentes que interactúan, así como también, es posible advertir la gran influencia que poseen los aspectos supra-personales en los agentes individuales, tal como la interacción social que es autónoma.

Entonces, con lo visto hasta aquí, se puede considerar claramente que el debate sostenido entre las teorías de la atribución mental y la cognición corporizada para explicar la cognición social constituye un choque de paradigmas que ha dado paso a un campo de investigación teórico-empírico muy fértil en su área.

En este contexto, cabe señalar que los teóricos corporizados y enactivos aceptan que en ocasiones recurramos a prácticas de atribución mental, pero sólo cuando nuestras prácticas corporizadas fallan, es decir, en encuentros extraños y de conductas inexplicables. Asimismo, las teorías de la atribución mental no están comprometidas con que la adscripción mental sea la *única* forma de interacción social, lo que defienden es que ésta es la *principal* forma de navegar exitosamente en el mundo social.

Como vemos, ambos enfoques parten de supuestos divergentes entre sí y cada uno defiende explicaciones en términos casi excluyentes, lo cual, considero, obstaculiza la posibilidad de ampliar sus posturas y, por ende, la investigación. Por lo que, en su lugar, defiendo que un marco integrativo e incluyente es más fructífero, pues en realidad ambos marcos conceptuales explican aspectos distintos, por tanto, ambos contribuyen con

aportaciones significativas al tema de interés. A esto vuelvo más adelante, por ahora me interesa resaltar las consecuencias que implica esta sección.

1.4.3. Consecuencias

1. Los computacionalistas explican la cognición social mediante el proceso de atribución mental, mientras que los enactivos priman la interacción como el recurso explicativo de la misma, además de poner en primer plano la percepción directa del otro. De ahí, que se les etiquete como “atribucionistas” a los primeros y como “interaccionistas” a los segundos.
2. Los atribucionistas defienden una perspectiva epistémica de 3ª persona, mientras que los interaccionistas defienden una postura epistémica de 2ª persona.
3. En consecuencia, los atribucionistas otorgan un papel de *observador* a los interactuantes, mientras que los interaccionistas les otorgan un papel de *participantes* involucrados o activos.
4. Los atribucionistas explican la cognición social a partir de los procesos que ocurren en la cabeza del agente individual, mientras que los interaccionistas lo hacen por enfatizar el aspecto relacional de la interacción intercorporal.

De ello, el proceso de cada postura puede instanciarse de la siguiente manera:

Proceso de la cognición social computacionalista:

El sujeto adscriptor *observa* al otro y percibe la información relevante (señales del entorno y conducta manifiesta) → procesa tal información mediante el proceso de atribución mental postulado (ya sea TT o TS o una hibridación de ellos), lo que le permite producir una representación del otro y de sus estados intencionales → obtiene un estado resultante, el cual atribuye al otro. Con ello, se explica y predice la conducta del otro. Respecto a la *acción* social, es la representación del otro y de sus estados intencionales lo que le permite al sujeto adscriptor planificar cómo actuar de una manera socialmente apropiada, ante lo cual, responde ejecutando ese plan con comandos motores lo mejor posible.

Proceso de la cognición social enactiva:

En una interacción diádica, ambos agentes se perciben directamente y actúan en consecuencia (hay un acoplamiento sensoriomotor entre ellos en donde sus acciones son guiadas perceptualmente), lo que resulta en que ambos participan en una mutualidad directa que da lugar a participar en la co-creación de nuevos significados propios de esa

interacción específica y por tanto ambos responden a exigencias concretas de la situación. Con ello, ambos interactuantes participan y responden a la mutualidad que la interacción en sí misma exige.

Tras esta breve contrastación atribucionistas vs corporizados, ahora, corresponde pasar al tema de la cognición moral, en donde abordo los principales debates de la psicología del desarrollo moral por una parte y, por otra, presento los acercamientos enactivos al estudio de la moral. Esto permite contextualizar de dónde provienen los elementos de mi propuesta integrativa.

1.5. Cognición moral

En esta sección me enfoco, primero, en presentar los debates relevantes de la psicología del desarrollo moral debido a que sus avances y aportaciones han sustanciado el progreso del estudio de la cognición moral. En ello, se verá que en los años 2000 surgió un giro metodológico en la disciplina, lo que dio lugar a considerar distintos aspectos de la moral y distintas vías para abordar su estudio. Esto ha resultado en propuestas de nuevos paradigmas, entre ellos, “el nuevo paradigma de una moral innata”.

En segundo lugar, presento las aproximaciones del enactivismo respecto al estudio de la moral que, si bien, es un paradigma relativamente nuevo, también ha presentado un significativo progreso en la materia. A este respecto, es importante destacar que el enactivismo busca *complementar* y no *sustituir* las explicaciones cognitivas morales disponibles, esto porque intenta capturar sus dimensiones fenomenológicas, afectivas, interactivas y corporizadas, aspectos que no han sido abordados (en toda regla) en las explicaciones tradicionales y por tanto es plausible proponerlos como objetos de estudio que complementan la explicación moral.

1.5.1. Psicología del desarrollo moral

La inquietud por comprender qué caracteriza la moralidad humana, cómo opera, qué procesos son centrales a ella y cuál es su naturaleza, ha motivado un excepcional trabajo multidisciplinario que resulta en diferentes explicaciones teóricas-empíricas sobre nuestras competencias morales.

En principio, en la psicología del desarrollo se buscaba explicar la naturaleza de los juicios morales, debido a que cuando se trata de moralidad, la cuestión más básica refiere a nuestra capacidad de orientación normativa, i.e. a nuestra capacidad de estar motivados

por normas de comportamiento a través de juicios sobre cómo las personas deben actuar y responder en diversas circunstancias moralmente relevantes (Doris *et.al.*, 2020).

Al respecto, se ha articulado un álgido debate en torno a explicar la naturaleza del juicio moral, y su forma más general concierne a propuestas de corte racionalista, sentimentalista, intuicionista y de procesamiento dual. Dichas posturas intentan ofrecer explicaciones cuestionando si el fundamento habilitador del juicio moral es emocional, racional, intuitivo, o bien, si se trata de procesamientos de un sistema cognitivo dual en el cerebro.

En sus inicios, la investigación psicológica del desarrollo moral estuvo dominada por propuestas de corte racionalista, siendo la más aceptada la “Teoría de las etapas” de Kohlberg (1969), cuyo enfoque se centraba en el proceso de maduración del razonamiento moral, las orientaciones y los principios morales del sujeto. Kohlberg, basado en el paradigma piagetiano (Piaget, 1965), afirmaba que el razonamiento moral es desarrollado mediante una secuencia progresiva, estricta e invariable de etapas en las que el infante mejora sus habilidades de razonamiento.

Este enfoque plantea que el juicio moral es causado por un proceso de razonamiento y reflexión, cuya metodología se limitaba a considerar los juicios automáticos de niños y adolescentes (varones) a partir de dilemas morales problemáticos, en los que un individuo se encuentra en una situación tal que, sin importar lo que decida hacer, entrará en conflicto con alguna norma. Este modelo estrechó el alcance del estudio de la cognición moral, socavando el papel de otros procesos que pudieran desempeñar algún rol en el dominio moral (por ejemplo: emociones, comprensión de intenciones, influencia del entorno, etc.).

No obstante, en los últimos años, los descubrimientos acerca de intuiciones, emociones, la forma en la que los mecanismos automáticos interactúan con procesos racionales, entre otros, ha llevado a avances metodológicos que son el comienzo de una nueva era en el estudio de la moralidad. Así, tal giro metodológico en la investigación psicológica ha motivado a realizar estudios multidisciplinarios que intentan integrar todos los ángulos posibles de abordaje del tema de interés, lo que ha encaminado a algunos investigadores a poner mayor énfasis en los procesos emocionales (Nichols 2004, Prinz 2007), o en la intuición (Haidt 2001, Hauser 2006, Mikhail 2007), o en la noción de daño como piedra angular en el estudio de la cognición moral (Gray, 2018), entre otros. De este modo, se han desplegado nuevas áreas que tradicionalmente no habían sido conceptualizadas como parte del estudio del desarrollo moral.

En este contexto de revolución metodológica, surgen investigaciones que sugieren la existencia de capacidades proto-morales en edades muy tempranas, lo que conduce a la especulación de propuestas nativistas. No obstante, cabe mencionar que durante mucho tiempo los científicos del desarrollo fueron renuentes a la idea de que pudiera existir una facultad moral o un sentido de la justicia innato, pero, en la actualidad, un cuerpo importante de investigación dentro de la ciencia cognitiva está confluyendo en la existencia de un innatismo moral, dando paso con ello a la creación de “el nuevo paradigma de una moral innata” (Mikhail, 2008, 2011; Haidt y Joseph 2004, 2007; Hauser, 2006; Haidt 2012; Hamlin, 2013)¹³.

Bajo este paradigma, se ha intentado identificar, por ejemplo, cuáles son las bases psicológicas a través de las cuales las diferentes culturas crean la gran variedad de sistemas morales. Haidt y Joseph (2004, 2007) proponen cinco: 1) daño y cuidado, 2) justicia y equidad, 3) filiación y lealtad, 4) autoridad y respeto, y, 5) pureza y santidad.

En esta línea, también se ha sugerido que los humanos heredan genéticamente¹⁴ una "gramática moral" o una propensión a desarrollar reglas morales específicas (por ejemplo, prohibir el sexo con un hermano; Hauser, 2006); categorías de evaluación moral (por ejemplo, cuidado/daño, justicia/trampa, pureza/degradación; Haidt 2012); o una disposición a distinguir las normas morales de las convencionales (Turiel, 2002). Asimismo, existe la especulación acerca de la existencia de un núcleo moral innato (Hamlin, 2013; García-Ramírez, 2019) o de una propensión natural para el altruismo (Werneken y Tomassello, 2009)¹⁵.

Aunque los nativistas morales no presentan un frente unificado en sus propuestas, lo que comparten es que tienden a considerar que la naturaleza proporciona una base críticamente importante o un “primer borrador” del desarrollo moral, con lo que revelan un acercamiento

¹³ La existencia de una capacidad innata para la respuesta moral es sostenida de uno u otro modo por todos ellos y en esta línea también se encuentran: Greene, Pinker, Gazzaniga, Pizarro, Nichols, Dwers, Bloom, Cushman, Prinz, Doris y Carruthers.

¹⁴ Por supuesto, como se ha hecho ver, existen otras concepciones respecto a las formas de transmisión de la herencia. Otro ejemplo de ello se encuentra en la síntesis extendida (*evo-devo*).

¹⁵ Desde explicaciones evolucionistas de la moral, Darwin, Hamilton y Trivers son pioneros en sugerir un compuesto innato de la misma (Rosas, 2011, p. 209).

al modelo de la psicología evolutiva de la Escuela de Santa Bárbara (ver Heyes, 2019, pp.5-6).

El nuevo paradigma de una moral innata propone que la moralidad constituye un tipo de adaptación evolutiva y se explica como una capacidad emergente más que aprendida, aunque por supuesto requiera del moldeamiento del contexto y la cultura. Entonces, su principal postulado es que la moralidad se transmite mediante la ascendencia biológica y gracias a ella es posible la transmisión cultural. De ahí que otorguen un papel primario a la naturaleza.

Ahora bien, muchos comportamientos o principios como "ayudar" o "no dañar a otros" parecen constituir universales morales y estar presentes en edades muy tempranas. Hoy existe evidencia que nos da buenas razones para suponer que algunas capacidades morales o proto-morales pueden ser innatas y universales.

En general, los universales morales se han estudiado experimentalmente a partir del dilema del *tranvía*, un trabajo replicado con una muestra de más de 200,000 sujetos adultos en 120 países, obteniéndose sistemáticamente las mismas respuestas (Mikhail, 2008, 2011). Este dilema consiste en una situación en la cual un conductor de tren observa que cinco personas están en la vía, y debe elegir si atropellarlas o cambiar el tren hacia una vía adyacente en la que se encuentra una sola persona. La mayoría de los sujetos no tienen dudas en que lo correcto es cambiar la vía y salvar a la mayoría de las personas, aunque será inevitable que muera una persona. En un segundo ejemplo, se plantea la posibilidad de empujar a una persona obesa desde un puente con el objetivo de parar a un tren que va a atropellar a cinco personas. La respuesta en este caso suele ser: "No, no es moral empujar a alguien para salvar cinco vidas". Estas respuestas se obtienen en segundos, pero la cuestión relevante para los investigadores morales es ¿por qué sacrificar una vida en el primer ejemplo es aceptable, y en el segundo es inaceptable? Y la siguiente pregunta sería ¿por qué las personas sistemáticamente dicen sí, al primer dilema y no, al segundo? La respuesta es tan rápida que cabe preguntar ¿de dónde viene? ¿por qué todos parecen estar de acuerdo? (Quiroga-Méndez, 2013, p.182).

Como es de esperarse, debido a que las normas morales varían entre las culturas y el periodo histórico es común pensar que la moralidad es un aprendizaje que ocurre en la infancia a partir del contexto y, en esta línea, el *aprendizaje social* explicaría la diferencia moral entre individuos procedentes de distintas culturas, épocas, o sociedades. Pero ¿qué

es lo que explica las semejanzas de las respuestas morales? ¿cómo explicar las respuestas morales que aparecen de forma automática? ¿podemos hablar de universales morales constituyentes de la cognición moral? Arguyo que es plausible pensar que es mediante un conocimiento moral general que nos orienta normativamente hacia nociones universales, mismo que proviene, en parte, de nuestra herencia evolutiva; tema que desarrollo más adelante.

Ahora bien, hasta aquí he presentado el contexto de la revolución metodológica que surgió en el campo de la psicología del desarrollo moral, lo que permite considerar distintos marcos conceptuales que expliquen distintos aspectos del complejo espectro moral para arribar a una explicación integral de la misma. En particular, es en este contexto en el que se puede hablar de capacidades morales o proto-morales con carácter universal y tal vez innato.

Ahora, corresponde hacer explícitos los acercamientos enactivos al estudio de la cognición moral, lo que también es posible gracias a tal revolución metodológica.

1.5.2. Aproximaciones enactivas a la moral

En este punto es preciso tener en cuenta que las pretensiones del enactivismo son moderadas, pues intenta *complementar* las explicaciones ya existentes, mas no *inventar la rueda desde cero*. Hecha esta salvedad, comienzo su desarrollo.

La perspectiva enactiva toma como punto de partida su concepción de agencia e interacción y busca capturar las dimensiones fenomenológicas, afectivas, interactivas y corporizadas de la moral desde una metodología *relacional*.

El primer acercamiento enactivo a los temas morales se encuentra en el trabajo titulado *Ethical Know-how: action, wisdom, and cognition* de Francisco Varela (1999), quien concibe que la ética emerge de un confrontamiento corporizado y situado con el mundo, más que de procesos explícitos de reflexión y deliberación que siguen principios abstractos, resaltando con ello la acción pre-reflexiva en el campo de la moral.

Un segundo acercamiento es presentado por Colombetti y Torrance en 2009. Los autores proponen un cambio en la unidad de análisis de la ética: en lugar de enfocarse en el individuo y en la responsabilidad individual, sostienen la pertinencia de un enfoque en la interacción entre agentes y en los “matices morales” que emergen de ella. Esto plantea que el “carácter moral” o los “matices morales” que surgen de las interacciones sociales deben

de ser tomados en cuenta y, en consecuencia, el dominio de las interacciones y su dimensión *relacional* deben tener un lugar en el estudio de la moral.

En 2013, De Souza sigue la línea de estudio de la ética propuesta por Varela y explica las condiciones que dan lugar al desarrollo de un saber-cómo pre-reflexivo de la ética en relación con un “mundo ético”, que resulta en un tipo práctico de “experticia ética”.

Otra línea de acercamiento es representada por autores que han buscado los puntos de encuentro entre el enactivismo y la ética del cuidado (Loaiza, 2019; Urban, 2014, 2015; y van Grunsven, 2018), un enfoque derivado de las propuestas feministas de la ética. En breve, la ética del cuidado se enfoca en el aspecto relacional y afectivo de la misma; concibe a los agentes como vulnerables e interdependientes y, desde una ontología relacional, propone *el cuidado* como su valor central. Así, la convergencia con el enactivismo es que ambos enfoques parten de una metodología relacional; conciben a los agentes como intrínsecamente vulnerables y a las interacciones entre ellos como inherentemente afectivas y, por último, explican la ética a partir de la mutualidad presente en tales interacciones afectivas.

En 2021, Métais y Villalobos suman una contribución fenomenológica al enfoque interactivo de Colombetti y Torrance, poniéndolo en diálogo con el trabajo de Emmanuel Levinas y enfocándose en el aspecto corporizado y afectivo de la experiencia ética que éste propone.

En particular, Levinas sostiene que en *el encuentro con el otro* hay una dimensión fenomenológica de responsabilidad que se nos presenta como un llamado a responder, *i.e.*, en la interacción entre agentes emerge, a nivel experiencial, un llamado de responsabilidad hacia el otro y tal responsabilidad representa el deber moral Levinasiano. Así, Métais y Villalobos destacan la experiencia subjetiva de un deber moral fenomenológico presente en las interacciones sociales, en donde el agente percibe, en *el rostro del otro*, mandatos no-conceptuales y experimenta un enfático “cambio ético”: del egoísmo al altruismo. En otras palabras, *el yo pierde su primacía por el otro* y es llamado a sacrificarse, es decir, ante la alteridad, el agente experimenta un cambio fenomenológico respecto a sus involucramientos con el mundo: ahora es llamado a renunciar a su egoísmo y a responder de forma altruista, *i.e.*, a sacrificarse por el otro. Dicho “cambio ético” es la base de las relaciones éticas entre los agentes involucrados en las interacciones, según los autores.

Si bien, estos enfoques representan un gran avance en el desarrollo de una moral enactiva, tanto por destacar su aspecto interactivo y fenomenológico, como también, por poner de relieve la convergencia del enactivismo con la ética del cuidado, aún no se ha hecho explícito en qué consistiría una normatividad moral enactiva. No obstante, en 2022, mi colega Laura Mojica y yo presentamos una propuesta de la misma, la cual desarrollo en el capítulo 3.

Hasta aquí he presentado el contexto que enmarca las discusiones pertinentes a la cognición moral; por una parte, abordé la revolución metodológica que da lugar al nuevo paradigma de una moral innata en la ciencia cognitiva y por otra, expuse las aproximaciones enactivas al estudio de la ética y de la moral.

A partir de este panorama enriquecido por estudios multidisciplinarios se puede argüir que la cognición moral humana es dinámica y compleja, y este proyecto aspira entender los procesos cognitivos implicados en la misma, lo que permitirá responder puntualmente a la pregunta eje de esta investigación: ¿qué explica y qué resuelve empíricamente una que no lo haga la otra?

Con el objetivo de ofrecer una respuesta, dedico los siguientes dos capítulos (capítulos 2 y 3) a un análisis detallado de los argumentos que sostiene cada marco conceptual; el capítulo 2 dedicado puramente al Computacionalismo y el capítulo 3 puramente al Enactivismo. Por último, presento las consecuencias de este apartado que considero relevantes.

1.5.3. Consecuencias

1. El giro metodológico que se dio en la psicología del desarrollo autoriza explorar diversas posturas a fin de cumplir con el abordaje de los diversos elementos (morales y no morales) que constituyen la cognición moral.
2. El paradigma innato de la moral cuenta con suficiente apoyo empírico (el cual presento en la sección 2.5.7), que permite sugerir que los seres humanos poseen, al menos, ciertas nociones morales que son dotadas evolutivamente.
3. Las pretensiones teóricas del enactivismo respecto al estudio de la moral son moderadas, debido a que éste busca *complementar* y no *sustituir* las explicaciones morales disponibles, por lo que es plausible considerar los aspectos en los que se enfoca como objetos de estudio capaces de complementar las explicaciones tradicionales.

II. Preámbulo a los capítulos 2 y 3

Puntualmente, en el capítulo 2 abordo la evidencia empírica que permite argumentar a favor de: 1) la existencia de un tipo de conocimiento moral general que nos orienta normativamente, el cual (de manera parcial) es dotado evolutivamente y como tal permite sostener la objetividad moral, y, 2) develar los procesamientos funcionalistas-cognitivos que habilitan nuestras facultades morales, tales como el comportamiento moral y los juicios morales. Y, en el capítulo 3, abordo las explicaciones enactivas pertinentes a la cognición moral para argumentar: 1) a favor de una normatividad enactiva capaz de efectuar evaluaciones morales *situadas*¹⁶, 2) explicar los componentes corporizados que intervienen en la acción moral pre-reflexiva y en los juicios morales efectuados en 2ª persona, y, 3) para dar cuenta del papel que desempeñan los aspectos supra-personales (autónomos), tales como la interacción, el contexto y la cultura en la cognición moral desde una metodología *relacional*.

Tómese en cuenta que en cada capítulo abordo las explicaciones que ofrece el computacionalismo (cap. 2) y el enactivismo (cap. 3) respectivamente, lo que efectúo de manera *purista*, es decir, sin aludir a las aportaciones hechas por su respectiva contraparte teórica. Esta aclaración es relevante para evitar contradicciones a vista del lector.

Ahora bien, en este punto, considero útil precisar las implicaciones metodológicas que derivan de cada marco conceptual, así como también, sus entidades teóricas núcleo y sus principales niveles de análisis:

	Computacionalismo	Enactivismo
Modo-cognitivo	Computacionalista	Enactivo
Metodología	Individual/interna	Relacional

¹⁶ Se examina la inmediatez de la percepción y de la acción en una situación concreta, en donde la afirmación central es que la percepción de la situación provoca nuestras acciones (Varela, 1999).

Entidad teórica núcleo	Representación	Acción
Principal nivel de análisis	Sub-personal ¹⁷	Perceptual ¹⁸ y supra-personal ¹⁹

Se trata de abordar dos modos-cognitivos, dos metodologías y dos entidades núcleo que habilitan la cognición moral humana. Respecto a los modos-cognitivos asociados a cada perspectiva, en el capítulo 1 expuse detalladamente la naturaleza y los principios de cada uno de ellas, en donde se vio que el computacionalismo entiende la cognición como procesamientos de información representacionales que se rigen por el uso de reglas o principios, mientras que el enactivismo la entiende como acoplamientos sensoriomotores dinámicos agente-entorno que dan lugar a procesos de *hacer-sentido*. En breve, los dos

¹⁷ Es *sub-personal* en el sentido de que el sujeto nos es necesariamente consciente de ello, aunque, por supuesto, también se alude al nivel *personal* debido a que se habla de deseos, creencias, tomas de decisión, etc. Aquí es pertinente enfatizar que “en las ciencias cognitivas es un lugar común distinguir entre niveles de explicación *personal* y *subpersonal*. Esta distinción se remonta a Dennett (1969). La idea general es que el nivel *personal* se ocupa de las personas y su comportamiento e incluye vocabulario intencional (por ejemplo, “siente dolor/tiene la intención de abrir la ventana/está enojado” son descripciones del nivel *personal*), mientras que el nivel *subpersonal* se ocupa de los mecanismos cognitivos y neuronales que cumplen diferentes funciones en un sistema cognitivo y que pueden describirse en un vocabulario no intencional (por ejemplo, “las neuronas espejo se activan”). Tiene sentido decir, por ejemplo, “Me siento enojado” y “la corteza orbitofrontal lateral está activa”, pero no tiene sentido decir “mi cerebro se siente enojado” o “activé mi corteza orbitofrontal lateral” [...] La relación exacta entre los diferentes niveles es un tema muy debatido y no es posible dar una explicación detallada al respecto aquí (pero véase, por ejemplo, Colombo, 2012 para un informe reciente sobre el tema). Sin embargo, asumimos que las explicaciones en diferentes niveles deben considerarse mutuamente informativas.” (Bohl y van den Bos, 2012, pp. 4-5, nota al pie 11, las cursivas son mías).

¹⁸ Con fines de higiene conceptual es conveniente aclarar que la “percepción” en el computacionalismo es necesariamente representacional, mientras que en el enactivismo se entiende en una interdependencia entre el agente y el entorno, es decir, la acción es guiada perceptualmente porque se da en un mundo que es inseparable de nuestras capacidades sensoriomotoras.

¹⁹ El nivel *supra-personal* refiere a factores que se encuentran fuera de un individuo o no son reducibles a procesos internos de un individuo, como, por ejemplo, el entorno físico y social, el contexto, la dinámica temporal de la situación, etc. Al respecto, los interaccionistas han considerado que no es que los computacionalistas nieguen que los factores supraindividuales desempeñan un papel en la comprensión social o en la cognición, sino que, han minimizado su importancia en los mismos (ver Bohl y van den Bos, 2012, p. 3, nota al pie 5). En este sentido, Thompson (2007), por ejemplo, ha señalado que los procesos de interacción social no son reducibles a procesos en la estructura cognitiva interna del sujeto que interactúa, puesto que las interacciones implican *involucramientos físicos* con otras personas, lo cual no sólo es la *representación* de la persona (y de sus estados intencionales), sino que interactuamos con las personas de una manera que no podemos hacerlo con las representaciones; lo que deriva en la necesidad de sumar un análisis *relacional* del nivel supra-individual.

modos-cognitivos refieren a que uno es guiado por principios y ejecutado mediante procesamientos de información y el otro es guiado perceptualmente en un confrontamiento directo con el mundo, de lo que derivan consecuencias distintas.

Concerniente a las metodologías, corresponde hacer explícito cuál emplea cada marco conceptual, esto con la finalidad de revelar en qué aspectos se enfoca cada uno de ellos. El computacionalismo parte de lo que se podría llamar una “metodología individualista”, en el sentido de que su principal unidad de análisis es el individuo y como tal los fenómenos cognitivos los explica a partir de los procesamientos (representacionales) internos que ocurren en el sujeto individual. En contraste, el enactivismo emplea una “metodología relacional” y suele defender una epistemología interaccionista, lo que implica otorgar *autonomía* a los procesos que van más allá del agente individual (tales como la interacción, el contexto y la cultura), los cuales no son reducibles a procesos cognitivos individuales. Estas características permiten subrayar las claras diferencias metodológicas que prevalecen entre ellos.

Ahora bien, otorgar *autonomía* a los procesos supra-personales busca hacer justicia a la dimensión social de la moral desde un punto de vista *relacional*, lo que tiene al menos tres implicaciones: 1) explica el aspecto *situado* de la moral a partir de los significados compartidos que emergen de la interacción en marcha, los cuales, por su naturaleza relacional, son irreducibles al individuo; 2) apela a una co-constitución agente-entorno, o, agente-agente en el caso social, en el que ambos agentes son *participantes* activos de la interacción, lo que da lugar a una mutualidad co-regulada y a una *dependencia constitutiva*, y; 3) da cuenta de la normatividad pre-establecida que trasciende al agente cognitivo y sus interacciones, explicando el papel de un agente activo genuinamente empotrado en ella (los detalles de estas implicaciones se verán en el capítulo 3).

Tocante a la metodología individualista del computacionalismo, cabe resaltar que ésta implica cierto compromiso respecto a una “metodología interna”, la cual refiere a que, para explicar un fenómeno cognitivo *X*, es necesario conocer el funcionamiento interno del organismo, o bien, qué sucede en la estructura interna del organismo. A partir de Chomsky se ha establecido esta metodología interna, puesto que el autor, en oposición al “conductismo”, buscaba explicar los procesos cognitivos internos del individuo, las formas en que procesa e ingresa la información y cómo organiza su propio comportamiento (García-Ramírez, 2019, p. 266).

Por supuesto, con esta afirmación el computacionalismo no niega la amplia intervención que tienen el contexto y la cultura, ni excluye al entorno, a la sociedad o a las interacciones (y sus relaciones) en los fenómenos cognitivos a estudiar, de hecho, los reconoce como elementos que les desarrollan y complejizan. En concreto, los explica como parte de lo que procesan o computan representacionalmente los sujetos y en términos de sus acciones o toma de decisiones significativas y motivadas.

Así, pues, la discrepancia entre el computacionalismo y el enactivismo radica en cómo sus respectivas metodologías conceptualizan las intervenciones de los procesos supra-personales en los fenómenos cognitivos. El asunto aquí, no es que ellos se nieguen, sino cómo los conceptualiza cada enfoque:

- El computacionalismo los explica como parte de lo que procesa o computa representacionalmente el individuo, lo que resulta en toma de decisiones o acciones significativas y motivadas por parte del sujeto individual.
- El enactivismo les otorga autonomía, con lo que reconoce tanto la co-constitución agente-entorno y su papel normativo relacional que trasciende a los individuos y sus interacciones, como también, los significados compartidos que de ello surgen, lo que se logra, no por representarlos, sino por el aspecto *relacional* que emerge de los acoplamientos dinámicos agente-entorno o agente-agente, según sea el caso.

Entonces, bajo la línea computacionalista, se trata de explorar los procesamientos y mecanismos (representacionales) que habilitan los fenómenos cognitivos y, para mis propósitos, los procesamientos de información guiados por principios que subyacen a la capacidad socio-moral humana, mientras que, desde el enactivismo, se trata de revisar las intervenciones de los componentes corporizados y de los aspectos *relacionales* de la misma.

Ahora bien, respecto al tema de las “entidades teóricas núcleo” que habilitan a cada marco conceptual, la intención es hablar de dos objetos de estudio que habilitan la cognición socio-moral humana: la *representación* y la *acción*; la primera dedicada a los procesamientos representacionales cognitivos del individuo habilitados por el computacionalismo y, la segunda, dedicada a los acoplamientos sensoriomotores dinámicos y co-regulados que se dan entre el agente y el entorno (natural o social) en un confrontamiento inmediato con el mundo, de los que emergen una miríada de significados y normatividades compartidas, irreductibles a mecanismos o cogniciones individuales. Así, la arquitectura cognitiva PI que

propongo es capaz de soportar estos dos modos-cognitivos y sus metodologías, al mismo tiempo que respeta los principios de cada postura.

Pero ¿qué es una *representación* y qué es una *acción*? Respecto a la *representación*, recordemos que el computacionalismo comúnmente la entiende como un objeto mental intencional (dirigido a algo en el mundo) con contenido y con propiedades semánticas, por tanto, con condiciones de corrección, por lo que la representación es normativa²⁰. La condición mínima para que un objeto mental cuente como *representación* es que sea una estructura portadora de información.

Por su parte, la noción de *acción* se refiere a dichos acoplamientos sensoriomotores dinámicos que se dan entre el agente y el entorno, en donde la acción es guiada por la percepción y la interacción es co-constitutiva, lo que genera conductas que implican responsabilidades (Steiner y Stewart, 2009, p. 529).

Veamos a continuación el capítulo 2 dedicado al computacionalismo, para después, en el capítulo 3 ver con detalle los argumentos enactivos.

²⁰ Cabe mencionar que, por otro lado, está la noción de representación que aceptarían algunos puntos de vista de la Cognición corporizada, la cual se enfoca en el problema de explicar cómo las representaciones adquieren significado, sosteniendo que su significado se fundamenta en las capacidades de un agente para *sentir* el mundo y *actuar* en él (Singh, 2018, p. 136). Tales representaciones portarían información específica del cuerpo.

Capítulo 2

Computacionalismo

2.1. Introducción

Mis objetivos en este capítulo son dos: 1) respecto a lo cognitivo: explicar los procesamientos funcionalistas-cognitivos que habilitan nuestras facultades morales desde una arquitectura cognitiva altamente interactiva, capaz de involucrar de manera sustancial la información pragmática en sus procesamientos; la llamada arquitectura *súper-modular* (García-Ramírez, 2019), y; 2) respecto a lo normativo: argumentar a favor de la existencia de un tipo de conocimiento moral general (parcialmente) innato que funge como herramienta evaluativa objetiva capaz de orientarnos normativamente hacia nociones universales de lo correcto/incorrecto, bueno/malo y en general a distinguir objetivamente lo que es moral de lo que no lo es.

Concerniente a (1), me enfoco en hacer explícita la arquitectura *súper-modular* y a partir de ella explico los procesamientos funcionalistas involucrados en la acción moral y en los juicios morales, basándome principalmente en la evidencia empírica de infantes pre-lingüísticos arrojada por la psicología del desarrollo moral. Como se ha hecho ver, los procesamientos funcionalistas-cognitivos son, en esencia, procesamientos de información efectuados mediante el uso de representaciones y guiados por ciertos principios.

Concerniente a (2), se trata de postular la existencia de un tipo de conocimiento moral general (o cuerpo de información) que nos es dotado evolutivamente, al menos de manera parcial, el cual es capaz de asegurar la objetividad moral y, con ello, evita la apropiación cultural de los fundamentos morales. Sugiero que este tipo de conocimiento funciona como una estructura sistemática que otorga las nociones universales de la moral y, en consecuencia, permite explicar las generalidades de la misma y asegurar su objetividad.

Para lograr mis dos objetivos de este capítulo, recurro principalmente a una explicación de corte computacionalista que defiende la existencia de un “núcleo moral innato” (Hamlin, 2013; García-Ramírez, 2019), cuyo planteamiento es que la cognición moral está constituida por un aparato cognitivo dotado evolutivamente que cambia conforme al desarrollo, el cual incluye mecanismos cognitivos como habilidades de evaluación social, principios morales prelingüísticos, entendimiento intencional y emocional de los demás,

sensibilidad social y sintonía empática temprana. Desde edades muy tempranas el ser humano muestra aspectos de estos procesamientos cognitivos que favorecen el funcionamiento y desarrollo de la cognición moral. En la sección 2.5.7. presento la evidencia empírica que lo apuntala, lo que me permitirá habilitar mi propuesta.

Antes de continuar, es preciso hacer una aclaración de orden restrictivo. Empleo la explicación del núcleo moral innato (NMI en adelante) para habilitar la parte computacionalista de mi marco integrativo, debido a que a partir de sus argumentos y evidencia empírica extraigo la propuesta del conocimiento moral innato y defino los procesamientos funcionalistas-cognitivos de la cognición socio-moral, éstos facultados por la arquitectura *súper-modular*. No obstante, se debe tener en cuenta que me separo de tal propuesta en el sentido de que, en mi marco integrativo, involucro dos metodologías (la individual y la relacional) y, en virtud de la metodología relacional, otorgo “autonomía” a los procesos supra-personales. Con ello, el marco integrativo que propongo para explicar la cognición socio-moral humana se enfoca tanto en el funcionamiento cognitivo (representacional) del sujeto, como también, en la autonomía de los procesos supra-personales y sus intervenciones; ambos como aspectos constitutivos de la moral.

La justificación de mi propuesta se basa al menos en tres razones: 1) en coherencia con la crítica interaccionista respecto a la cognición social que se vio en el capítulo 1, asumo que la cognición socio-moral no sólo debe estudiarse desde lo que ocurre representacionalmente en el *individuo*, sino también, desde lo que ocurre *entre individuos* y los significados que de ello surge desde un aspecto relacional, pues la moral es principalmente un fenómeno que se da *entre* agentes, y por tanto también debe estudiarse en contextos pragmáticos *relacionales* que la conciban como una estructura social; 2) siguiendo la tradición sociológica, el planteamiento es que las estructuras socialmente construidas, aunque son almacenadas y procesadas en sujetos individuales, no tienen que ver sólo con la cognición individual, sino que ellas se componen de patrones duraderos de significado que emergen socialmente, los cuales adquieren vida propia o independencia por trascender a los sujetos individuales; aspecto relacional que debe ser reconocido y estudiado, y; 3) como mencioné en el capítulo 1, en estudios evolutivos recientes se ha mostrado una genuina co-constitución agente-entorno por el alto impacto que tiene la cultura en las diferencias genéticas, por lo que, dar *autonomía* a los procesos supra-personales, no sólo es una concesión teórica, sino que es urgente para dar cuenta de la moral, o al menos así lo defenderé.

Ahora bien, el presente capítulo lo divido en cuatro apartados y su estructura es la siguiente: 1) Evolución, 2) Procesamiento funcionalista-cognitivo, 3) Cognición social computacionalista, 4) Cognición moral computacionalista.

En la sección de “Evolución” se verá que las explicaciones que apelan a la selección natural como la principal fuerza evolutiva para la emergencia de la moral humana, en realidad son limitadas, pues lo que logran explicar es la herencia biológica concerniente a la presencia y evolución de los instintos sociales (filial y parental) y de la empatía, mismos que representan la génesis de la moral. No obstante, para explicar el fenómeno completo es necesario recurrir al entrelazamiento de lo natural y lo cultural debido a que, para hablar de un sentido moral en toda regla, es preciso hablar de capacidades más sofisticadas tales como el razonamiento y el lenguaje, capacidades que evolucionaron en virtud de la sincronía de procesos evolutivos biológicos y procesos evolutivos culturales y no sólo provenientes de la selección natural como único mecanismo de herencia²¹. Así, sostengo que, al menos, el conocimiento moral general dotado evolutivamente es el producto de una interacción dinámica co-evolutiva bio-cultural.

Entonces, en la sección de “Evolución”, me enfoco en las contribuciones que tuvo la selección natural en el constructo de la moralidad con dos propósitos: a) para argumentar a favor de las predisposiciones naturales que tenemos para desarrollar tal conocimiento (i.e., como algo intrínseco al humano), y, 2) para aproximarme a una descripción de la disposición natural que posee el humano para desarrollar la moralidad.

En la sección de “Procesamiento funcionalista-cognitivo” me enfoco en detallar tal procesamiento y en los supuestos cognitivos que sostiene. En esta sección, abordo los compromisos funcionalistas-cognitivos que habilitan el modo de la cognición moral computacionalista.

En la sección de “Cognición social computacionalista” reviso brevemente las explicaciones socio-cognitivas innatistas-modulares debido a que son el precedente que permite ver los problemas que han dejado los debates tradicionales en torno a las concepciones nativistas de las mismas. Su análisis permite enfatizar qué es lo que debe intentar resolverse al momento de hablar de capacidades innatas.

²¹ Por supuesto, no es que los procesos culturales se nieguen, sino que estas perspectivas le otorgan un papel relegado en sus explicaciones, mismo que, considero, debe ser reivindicado.

Por último, en la sección de “Cognición moral computacionalista” efectúo dos tareas: 1) hago explícita la arquitectura *súper-modular* (capaz de superar aquellos debates tradicionales) y a partir de ello expongo detalladamente los argumentos y la evidencia empírica que sostiene y habilita al NMI, y; 2) presento sus implicaciones normativas. Esto constituiría la parte computacionalista de mi propuesta integrativa.

2.2. Evolución

¿Qué es bueno y qué es malo? ¿qué es correcto y qué es incorrecto? ¿contamos con una base objetiva que nos permite distinguirlos? ¿contamos con ciertas predisposiciones naturales que nos orientan normativamente para discernirlos? Y si es así, entonces ¿poseemos alguna predisposición natural acerca de la moral? O en su lugar ¿la moral es completamente adquirida y aprendida? En esta sección argumento que los seres humanos llegamos instruidos en algún grado con ciertas predisposiciones naturales que contribuyen a la conformación de un tipo de conocimiento moral general favorecido por la evolución.

Pero si es así, entonces cabe preguntar ¿cuál es el origen de tal conocimiento? ¿ese conocimiento es capaz de justificar por qué todas las poblaciones humanas tienden a desarrollar nociones morales? Para responder a ello, a continuación, abordo lo que considero un “problema de origen”.

2.2.1. Problema de origen

Como punto de partida considero un “problema de origen” respecto a la moral, el cual, sostengo, es posible abordarlo desde el argumento del conocimiento moral general dotado evolutivamente. Son dos las preguntas centrales de lo que llamo “el problema de origen”: 1) la primera es atinente a la *universalidad* innegable de la moral, es decir, parece que toda comunidad humana está destinada a desarrollar, al menos, algún tipo de moralidad. Si esto es así, la cuestión a responder sería ¿por qué? y; 2) si se acepta tal *universalidad* ¿qué constituiría dicho carácter universal?

Concerniente a (1), mi respuesta alude a un argumento nativista aceptado por algunos teóricos cognitivos, el cual refiere a que es posible que la moral sea parcialmente dotada por la evolución²², lo que justificaría su universalidad, pues como tal, sería parte natural de nuestro aparato cognitivo humano (Hamlin, 2013, 2015; García-Ramírez, 2019). Al

²² Tal parcialidad refiere a la base innata que se desarrolla y complejiza mediante la interacción con el entorno y con los otros.

respecto, se considera que en la actualidad existe apoyo empírico suficiente que nos da buenas razones para sugerirlo. Si esto es así, se podría decir que la moral es intrínseca al ser humano, por lo que es natural que toda comunidad humana tienda a desarrollarla (al menos, nociones de ella). Pero ¿qué constituye esa universalidad? Esto nos conduce a la cuestión (2).

Concerniente a (2), si se acepta tal universalidad ¿qué constituiría dicho carácter universal? Al parecer, en términos generales, la moral se constituye de un núcleo de orientaciones normativas que nos guía hacia nociones de lo correcto y lo incorrecto, es decir, la moral funge como una herramienta evaluativa capaz de guiar nuestros cursos de acción y es apta para hacernos discernir entre ellos bajo estructuras sistemáticas y generalizadas que configuran lo que consideramos como correcto/incorrecto, bueno/malo, aceptable/inaceptable, encomiable/despreciable, etc. En este contexto el carácter universal de la moral radicaría en admitir que las diversas comunidades humanas comparten, al menos, ciertas nociones amplias de esa estructura general y sistemática, la cual estaría compuesta de principios universalmente compartidos tales como “ayudar a otros es bueno”, “el sufrimiento es malo”, etc., principios que parecen ser innegablemente universales y que, en un sentido general, subyacen a nuestros juicios y cursos de acción moralmente relevantes.

En particular, existen aspectos notablemente universales de la estructura moral que permiten sugerir que existe una base común capaz de asegurar y explicar su objetividad²³, pues de lo contrario ¿cómo es posible que toda comunidad humana tienda a desarrollar sistemas morales o, al menos, nociones morales? o bien, ¿cómo podríamos explicar que las poblaciones humanas a lo largo de la historia (o gran parte de ella) han presentado rasgos o propensiones morales? Ya sea óptima o no, pero la moral parece ser un rasgo característico de las poblaciones humanas.

Y, asimismo ¿por qué parece que la moral apunta a una estructura subyacente similar de lo que se categoriza como correcto/incorrecto en general? Es decir, parece que hay una estructura sistemática y generalizada que habilita y guía lo que configura esas clasificaciones comunes de lo correcto/incorrecto, bueno/malo, lo que ha permitido a la

²³ Desde este punto de vista se sugiere que “hay verdades morales [...] y nadie inventó la moralidad, incluso nuestros antepasados remotos ya la tenían [...]. La moralidad se transmite a través de la descendencia biológica; la transmisión cultural es posible gracias a ella.” (García-Ramírez, 2019, en borrador, p. 282, nota al pie).

humanidad contar con nociones amplias y compartidas que subyacen al carácter universal de la moral²⁴; tales como los principios incontrovertiblemente aceptados de “ayudar a otros es bueno” y/o “el sufrimiento es malo”.

En este punto, se podría preguntar, ¿habría alguna señal de cuándo tuvo lugar esto en el orden filogenético? pese a que no es posible fechar este hecho, más adelante se verá que la simpatía, los instintos sociales y la intensa repetición de las estrategias que favorecieron la *aptitud* humana (*fitness*), contribuyeron decisivamente a la selección de los mecanismos que subyacen a nuestras nociones morales. Asimismo, se verá que la auto-domesticación del hombre, que tuvo lugar a partir de la vida en grupo, derivó en una cascada de características físicas y psicológicas que favorecieron la evolución de diversos aspectos, tales como la cooperación y la moralidad humana²⁵.

Entonces, si mis consideraciones son correctas, desde este punto de vista, parece que la moralidad es intrínseca al ser humano y su carácter universal refiere a la estructura sistemática y generalizada de nociones normativas que se comparten entre las diversas culturas.

Así, pienso que estas cuestiones favorecen el argumento de la existencia de un tipo de conocimiento moral general como parte de nuestra naturaleza humana, es decir, el planteamiento es considerar que la moral es parte de nuestro aparato cognitivo. Si esto es

²⁴ El planteamiento es análogo a la afirmación de que los seres humanos tienen mecanismos innatos de aprendizaje del lenguaje, lo cual no implica que el japonés o el italiano sean innatos en sí mismos; estamos preparados para aprender un idioma u otro, pero el entorno social determina cuál. Del mismo modo, es posible que la selección natural se haya interesado por el contenido de la moralidad, quizá favoreciendo los universales amplios y generales (ver Joyce, 2006, p. 10).

²⁵ En sentido concomitante, Mithen (1998) señala que: “En la mente apareció una fluidez cognitiva que no correspondía a un nuevo poder procesador, sino que reflejaba nuevas conexiones mentales. Y, por consiguiente, esta transformación mental tuvo lugar sin aumento del tamaño del cerebro. Supuso, básicamente, el origen de la capacidad simbólica exclusiva de la mente humana, con múltiples consecuencias para el comportamiento cazador-recolector [...]” (Mithen, 1998, p.223), en donde “[p]arece que el primer paso hacia la fluidez cognitiva fue una integración entre la inteligencia social y la inteligencia de la historia natural, algo que evidencian los primeros humanos modernos del Próximo Oriente, hace 100.000 años. [...] El último paso hacia una total fluidez cognitiva tuvo lugar en momentos ligeramente distintos en diferentes poblaciones hace entre 60.000 y 30.000 años, y conllevó una integración de la inteligencia técnica, provocando los cambios de comportamiento que se observan en lo que se llama el periodo de transición del Paleolítico Medio al Superior. En otras palabras, se creó una explosión cultural: la aparición de la mente moderna” (Mithen, 1998, pp. 207-208).

así, ahora el desafío es ofrecer una explicación evolutiva de los factores que pudieron dar lugar a ello, tema al que dedico la siguiente sección.

2.2.2. Psicología evolutiva: escuela de Santa Bárbara

Primero, respecto a las escuelas de la psicología evolucionista vistas, los teóricos suelen relacionar el computacionalismo con los principios de la escuela de Santa Bárbara, aunque por supuesto, tal relación depende de la versión computacionalista de la que se esté hablando²⁶.

Dados los principios de la versión computacionalista que aquí abordo, asumo tal vínculo, aunque en un sentido moderado, por lo que, en este apartado, el principal compromiso ontológico de fuerza selectiva como mecanismo de herencia es la “selección natural”. Así, el argumento de esta sección está comprometido fundamentalmente con el *adaptacionismo*, y el *modularismo*. En este contexto realizo el siguiente análisis.

Los principios de la psicología evolucionista de la escuela de Santa Bárbara tienen como objeto general tratar la moralidad como un conjunto de disposiciones y comportamientos adaptativos que evolucionaron a partir de disposiciones y comportamientos "prosociales" o "proto-morales" que ya se encontraban en nuestros ancestros humanos extintos. Basados en "La descendencia del hombre" de Darwin (1871), quien encontró precursores instintivos del sentido moral en aves y mamíferos, se establece este curso de investigación.

Asimismo, bajo esta aproximación darwinista, las explicaciones evolutivas del altruismo y de la cooperación humana ya han sugerido que "la biología podría eventualmente construir una explicación científica plausible de un núcleo de la moralidad humana" (Rosas, 2011, p. 209). En esta línea, un principal argumento es que la moralidad se opone (con frecuencia) al interés propio inmediato, y su plataforma psicológica implicaría la comprensión y la preocupación por el otro (Wilson en Joyce, 2018, p. 302). Así, se ha asumido que la moral debe ser dirigida a *otro* u *otros*, más que a uno mismo²⁷.

²⁶ La conexión entre la escuela de Santa Bárbara y el computacionalismo tiene historia y cuenta tanto con adeptos fervientes como Pinker, como también con otros autores que no se convencieron por completo de esta conexión (Perconti y Plebe, 2020).

²⁷ “Durante décadas, los investigadores evolucionistas, desde Richard Dawkins hasta Steven Pinker, han argumentado que los seres humanos son capaces de organizarse y cooperar de manera tan eficaz porque nuestra psicología ha sido moldeada por las fuerzas evolutivas de la selección de parentesco y del altruismo recíproco (reciprocidad)”. (Henrich, 2016, p. 142).

A la luz de ello, la motivación en esta sección es explorar la posibilidad de que la evolución favoreció predisposiciones naturales para desarrollar un tipo de conocimiento moral general capaz de orientarnos normativamente.

2.2.3. Desarrollo del argumento evolutivo

Se trata de hablar de dos tipos de herencias, la genética y la ecológica, la primera enfocada en los rasgos heredables y la segunda en las modificaciones ambientales. En esta sección me ciño a la herencia genética, es decir, a aquellos rasgos biológicos heredables que pudieron contribuir a la evolución de la moral, las cuales son al menos dos: 1) la condición de vulnerabilidad del cuerpo humano que exigió la implementación de estrategias cooperativas, y, 2) la presencia de empatía y de instintos sociales, paternos y filiales, los cuales, integrados con la intensa repetición del cúmulo de estrategias exitosas y perniciosas de nuestros antepasados, favorecieron la selección de mecanismos que conforman nuestras orientaciones normativas generales y universalmente compartidas. Esto constituye un núcleo biológico de rasgos heredables que componen la plataforma psicológica de la moral dirigida al bienestar y a la preocupación por los demás, o al menos así lo defenderé.

Es importante tener en cuenta que mi intención es únicamente aproximarme a una descripción sobre la naturaleza de ello a partir de un análisis de *grano grueso*, pues un escrutinio profundo de las explicaciones evolutivas disponibles sale del alcance de la presente tesis. Con esto en mente, comienzo el desarrollo del argumento.

En primer lugar, la hipótesis apunta a distinguir la evolución biológica y la evolución cultural, lo que permitirá sostener que el fenómeno de la moralidad surge de una interacción dinámica de procesos evolutivos biológicos y procesos evolutivos culturales, o bien, de lo que se llama la co-evolución de la naturaleza y de la cultura, pues alcanzar la sofisticación que implica la moral humana requiere reconocer tal entrelazamiento.

En este punto, cabe aclarar que al hablar de las dos herencias (biológica y ecológica) no significa que los defensores de una postura nieguen la fuerza evolutiva de la otra, pues en la actualidad se ha descartado una estricta separación de las mismas. Se trata más bien de

resaltar el papel que se le otorga a cada una de ellas, lo que permite analizar sus contribuciones al tema de interés²⁸.

Por el lado de lo biológico o natural, Darwin basa el sentido moral en los instintos sociales (que incluyen sentimiento filial y parental) y, en particular, en la simpatía. Para Darwin hay una reducción en número, especialización y fuerza de tales instintos en el hombre, lo cual posibilitó la base natural para el desarrollo moral, pero la evolución cultural fue la que permitió el avance de las cualidades morales. En este sentido, Darwin sostiene que la moral es un fenómeno de la historia cultural con raíces en la historia natural (Engels en Etzelmüller, 2016).

En esta sección la pregunta concreta es ¿en qué contribuyó la selección natural en la evolución de la moral? Según Darwin, la naturaleza nos dotó de instintos sociales que incluyen afectos paternos y filiales, sociabilidad y simpatía, lo que permitió al hombre vivir en grupo y con ello desarrollar el placer por los demás, la fidelidad, la voluntad de ayudar, etc. De esta manera, la facultad de vivir en grupo resultó en la auto-domesticación del hombre, lo que representó el inicio de una cascada de características físicas y psicológicas que, entrelazadas con procesos culturales, favorecieron la evolución de diversos aspectos, tales como la cooperación y la moralidad²⁹.

Revisemos el argumento darwinista respecto a cómo la selección natural favoreció la evolución de tales instintos y afectos sociales.

Darwin (1877) comienza por explorar dos marcos conceptuales, por un lado, el marco evolutivo del hombre como especie, y por otro, la ética normativa. El punto de partida de sus reflexiones es el supuesto de que el hombre primitivo, antecesor del hombre civilizado, ya poseía dichos instintos sociales que incluyen afectos paternos y filiales, sociabilidad y simpatía. El instinto de simpatía se considera la raíz de nuestro sentido moral debido a que

²⁸ Al respecto, “[h]ace tiempo que los científicos han descartado una estricta división entre ambos ----entre «naturaleza» y «crianza». Todo comportamiento tiene que estar influido en parte por la estructura genética del animal y en parte por el medio en que se desarrolla. Sin embargo, el peso relativo de una u otro varía considerablemente entre las especies, sobre todo por lo que se refiere a los distintos aspectos del comportamiento de una sola especie” (Mithen, 1998, p.214).

²⁹ En el capítulo 3 se verá que los procesos culturales constituyen un nicho cultural específico del humano, el cual encuentra su origen en la simpatía, lo que favoreció las prácticas de aloparentalidad, mismas que dieron lugar a las capacidades socio-morales en dos sentidos: 1) por favorecer la prosocialidad y 2) por beneficiar la simpatía directa con las expresiones de los sentimientos de otros.

está dirigido hacia el bienestar de los otros y de la comunidad, lo que contribuye a la plataforma psicológica que implica la comprensión y la preocupación por el otro.

De esta manera, en el transcurso de la evolución humana, desde los progenitores semi-humanos hasta los humanos modernos, hemos heredado instintos y, más tarde, impulsos instintivos dirigidos al bienestar de los demás, los cuales son considerados condición necesaria para la formación de nuestro sentido moral.

Inicialmente, el hombre no estaba interesado en preservar a la especie sino en preservar a su propia comunidad y tribu, para lo cual, tuvo que implementar estrategias de cooperación a fin de luchar por la supervivencia del grupo; estrategias cooperativas propiciadas, en origen, por aquellos instintos sociales. En este contexto, es importante señalar que, según Darwin, existe una notable correlación entre nuestra condición corporal humana con su específica vulnerabilidad, y la extraordinaria cooperación que hemos desarrollado como especie. El autor afirma que "el hombre es una de las criaturas más indefensas y menos defensivas del mundo", no obstante, parece que es precisamente este hecho el que contribuyó al éxito humano en la evolución.

Darwin señala que "no sabemos de qué especie, fuerte o débil, desciende el hombre, pero pudo haber sido una inmensa ventaja para el hombre haber surgido de alguna criatura comparativamente débil, puesto que, debido a nuestra vulnerabilidad, pudimos volvernos sociales y así adquirir nuestras cualidades mentales y sociales superiores" (en Engels, 2016).

El planteamiento es que, debido a la vulnerabilidad natural de la condición corporal humana (entre otros factores), el medio exigió al hombre implementar estrategias de supervivencia para lograr su preservación; para defenderse de ataques de exo-grupos, de desafíos o de desastres naturales, de amenazas potenciales del entorno etc., lo que logró mediante la cooperación entre miembros de una misma comunidad. Así, se considera que la vulnerabilidad natural del hombre y sus instintos y afectos sociales, incluida la empatía, fueron precedentes naturales que contribuyeron decisivamente a que la cooperación se convirtiera en una estrategia de lucha por la existencia, pues aseguró la supervivencia de las comunidades, además de otros beneficios que no hubiesen sido posibles de alcanzar de manera individual.

Ahora bien, en comparación con los primeros humanos, nuestros instintos se redujeron de varias maneras, en su número, su especialización y su fuerza. Según Darwin, la condición para el desarrollo de una moral genuina es precisamente esta reducción de los instintos junto con la evolución de la razón, del juicio y del lenguaje (Darwin [1877] 1989)³⁰.

En sentido concomitante, la reducción de los instintos y la disminución de la agresión humana (Pinker, 2011) es consistente con algunos rasgos asociados con la domesticación, tales como el encogimiento del cerebro humano, el adelgazamiento del cráneo y una reducción del dimorfismo sexual (Chen y Moyzis, 2018, p. 769; ver también Brüne, 2007; Jones, 2000)³¹.

Así, pues, aunque el primer fundamento u origen del sentido moral reside en los instintos sociales, incluida la simpatía como piedra angular (Darwin [1877] 1989), ellos no bastan por sí solos para explicar el fenómeno de la moral. Darwin afirma que la moralidad genuina consiste en el sentido moral, en un sentido del bien y del mal, siendo esto algo que sólo el humano posee.

Pero ¿de dónde provendría el sentido del bien y del mal? en breve, se ha argumentado que, al menos una noción del “bien”, entendida en sentido amplio y general, proviene de todas aquellas estrategias que beneficiaron la reproducción y la supervivencia (*fitness*) de nuestros antepasados, lo que favoreció la selección de ciertos mecanismos que contribuyen a nuestras evaluaciones universales de lo que valoramos y consideramos como “bueno”. En contraste, un sentido del “mal” resultó de todo aquello que implicó daño y fue nocivo para la preservación de la especie. Así, se puede pensar, en sentido amplio, que la selección de los mecanismos que favorecen nuestras nociones universales de lo bueno/malo evolucionaron en nuestros aparatos cognitivos a partir de aquellas estrategias

³⁰ En palabras de Darwin (citado en Engels, 2016): “[...] en comparación con nuestros primeros progenitores simioscos y humanos, nuestros instintos se reducen en varios aspectos, en número, especialización y fuerza. La condición para el desarrollo de una moral genuina es esta reducción de los instintos junto con la evolución de la razón, el juicio y el lenguaje, con “libre inteligencia” y “libre albedrío” (Darwin [1877] 1989, 71 y sigs.). Sin embargo, los instintos sociales todavía dan el impulso a nuestras acciones sociales y morales. No obstante, tienen que estar orientados por la razón.” (Engels en Etzelmüller, 2016, p. 239).

³¹ Por ejemplo, al respecto, Brüne dice: “Desde una perspectiva biológica, la mayor disputa con respecto a los cambios físicos en humanos anatómicamente modernos similares a la domesticación se refiere a una disminución leve pero mensurable del volumen cerebral de alrededor de 1.400 cm³ a aproximadamente 1.300 cm³, lo que podría interpretarse como un mayor apoyo a la hipótesis de la auto-domesticación humana” (Brüne, 2007, pág. 5).

benéficas o perjudiciales, al menos, a manera de orientaciones generales exitosas o perniciosas.

Ahora bien, según Darwin, una condición necesaria del sentido moral fue la evolución de capacidades intelectuales humanas como el razonamiento y el lenguaje verbal, y aquí es donde el poder del juicio moral marca la diferencia, pues, de acuerdo al autor, la capacidad para efectuarlos presupone razonamiento, lenguaje verbal y autoconsciencia, así como también, el entendimiento y la formulación de normas morales (ver Engels, 2016, en Etzelmüller, p.240). Pero ¿cuál es el origen de los juicios morales? Veamos lo que se ha considerado respecto a ello, al menos pertinente a su origen cognitivo.

- **Origen del juicio moral**

La naturaleza social y cooperativa del ser humano le ha permitido crear a lo largo de su historia sistemas cooperativos y de alianza fructíferos y exitosos, es decir, los humanos, al trabajar en conjunto, compartir recursos y ayudar a los necesitados, han logrado un progreso innegable que les permite estar más seguros, mejor alimentados, ser más fuertes, tener mejores estrategias, mejores herramientas, etc., lo que no habrían alcanzado de manera individual. Empero, estos sistemas cooperativos implican pagar por adelantado pequeños costos de cooperación para después, a largo plazo, estar mucho mejor (Hamlin, 2015, pp. 7-8).

Una consecuencia crucial de tales sistemas cooperativos es que son vulnerables a los tramposos (quienes no cooperan y obtienen beneficios del trabajo de otros). Así, esta vulnerabilidad requiere que los sujetos integrantes de los sistemas cooperativos sean capaces de detectar a los tramposos, lo que implica que sean capaces de saber las intenciones de los demás. De ello, parece que el origen cognitivo del juicio moral se halla (al menos en parte) en tal vulnerabilidad a los tramposos, pues parece necesario saber leer, inferir y predecir las intenciones y los estados mentales de los otros para detectarlos y juzgarlos como tal.

Entonces, el origen cognitivo de los juicios morales parece fundamentarse en las capacidades socio-cognitivas de *leer* otras mentes, capacidad que también implicó importantes sincronías de procesos bio-culturales para su evolución (una hipótesis de ello es la *hipótesis del cerebro social*, propuesta por Dunbar).

En torno a ello, se ha argumentado que la moralidad humana, impulsada por capacidades sociales, se ha desarrollado tanto filogenéticamente como ontogénicamente debido a que es el fundamento clave para que los sistemas cooperativos florezcan y prosperen. De ahí que se piense que la moral y la cooperación co-evolucionaron, y que tal co-evolución fue la que permitió la supervivencia y el éxito de la especie.

Ahora bien, Darwin describe la moralidad humana como un fenómeno complejo que involucra capacidades racionales y emocionales muy distintas, particularmente simpatía y virtudes. Las virtudes sociales las consideraba como una expresión de nuestro sentido moral y ahora son descritas como "la descendencia refinada y ampliada" de los instintos sociales de los animales y de los primeros humanos (Engels, 2016).

En torno a ello, Darwin reconocía que los instintos, las habilidades mentales y el lenguaje verbal evolucionaron a través de los mecanismos generales de la evolución, pero ya desde la primera evolución de la humanidad, en adelante, se entrelazaron con los hábitos y con factores culturales³² (ver Engels, 2016, p. 244).

En sentido estricto, Darwin defendía la teoría lamarckiana de la herencia de los caracteres adquiridos, con lo que concebía que los hábitos, si se practicaban durante un largo período de tiempo, podrían convertirse en rasgos de los organismos. De este modo y con el paso del tiempo, la moral se convirtió cada vez más en un fenómeno cultural con raíces en la historia natural (Engels, 2016, en Etzelmüller, p.244).

La perspectiva de Darwin sobre las "generaciones futuras" era optimista, porque sostenía que "podemos esperar que los hábitos virtuosos se fortalezcan, y que *tal vez* sean fijados por herencia" (Darwin [1877] 1989, 129 y sigs., citado en Engels, 2016, p.245).

No obstante, por otro lado, también afirmaba:

³² Engels (2016) señala esta afirmación de Darwin de la siguiente manera: "Las facultades morales de los humanos tienen sus raíces en la historia natural del hombre, en los instintos sociales así como en las habilidades mentales de los animales prehumanos. Pero estas raíces no son suficientes para explicar y describir nuestro sentido moral o conciencia. En el curso de la evolución humana, desde los predecesores animales, pasando por los progenitores semihumanos hasta el hombre moderno, hemos heredado instintos y, más tarde, impulsos instintivos, dirigidos al bienestar de los demás, que son la condición necesaria para la formación de nuestro sentido moral. "Un hombre que no poseyera ningún rastro de tales instintos sería un monstruo antinatural" (Darwin [1877] 1989, 116). Los instintos, las capacidades mentales y el lenguaje verbal como condiciones de la moral humana evolucionaron a través de los mecanismos generales de la evolución, pero ya desde la evolución temprana de la humanidad se entrelazaron con hábitos y factores culturales, un proceso que hoy se denomina "coevolución" de la naturaleza y la cultura" (Engels, 2016, p. 244).

"Por importante que haya sido la lucha por la existencia y aún lo sea, en lo que respecta a la parte más elevada de la naturaleza del hombre, existen otras agencias más importantes. Pues las cualidades morales avanzan, directa o indirectamente, mucho más por los efectos del hábito, las facultades de razonamiento, la instrucción, la religión, etc., que por la Selección Natural; aunque a esta última agencia pueden atribuirse con seguridad los instintos sociales, que proporcionaron la base para el desarrollo del sentido moral. (Darwin [1877] 1989, II, 643)". (citado en Engels, 2016, p. 245).

Así, Darwin concebía la moral como un fenómeno cultural pero enraizado en la historia natural, y justo ahí se sitúa mi propuesta.

Como se puede ver, distintos aspectos y capacidades que contribuyeron a la evolución de la moral surgieron de las fuertes interacciones entre lo natural y lo cultural, por lo que es plausible ir más allá de las suposiciones de la escuela de Santa Bárbara y buscar hipótesis que hagan justicia a dichas interacciones (mi mejor candidata para ello, es la *hipótesis del cerebro cultural*, la cual presento en el siguiente capítulo).

Por supuesto, las explicaciones evolucionistas de la moral han rebasado estos argumentos darwinistas, así como también, por su parte, las discusiones respecto a los posibles entrelazamientos entre lo biológico y lo cultural siguen constituyendo debates encendidos sobre el tema de interés.

Empero, mi intención aquí ha sido únicamente aproximarme a una descripción sobre la naturaleza de las predisposiciones morales del hombre, lo que, en última instancia, contribuye a explicar parcialmente la presencia del conocimiento moral general dotado evolutivamente. Esto último se verá más adelante, pues para ello es necesario dar cuenta de los procesamientos culturales.

Entonces, hasta aquí tenemos que dicha descripción natural apunta a resaltar al menos dos aspectos: 1) el papel que jugó la vulnerabilidad corporal del hombre en la evolución de las conductas cooperativas, y, 2) la presencia de instintos sociales y empatía en los primeros humanos, lo que contribuyó decisivamente a la plataforma psicológica de la moral, la cual se dirige al bienestar y a la preocupación por los otros.

Esto constituyó un núcleo biológico de rasgos heredables, que junto con la intensa repetición de las estrategias exitosas y/o perniciosas de nuestros antepasados, contribuyó a la selección de aquellos mecanismos que subyacen a nuestras evaluaciones generales de lo que consideramos como "bueno" y "malo".

Así, pues, si mis consideraciones son correctas, es plausible pensar que la herencia genética se corresponde bien con los aspectos mencionados, sin embargo, aún falta ver la herencia ecológica y con ello la contribución de los procesos culturales, tema que se verá en el capítulo 3.

Por el momento, subrayo tres consecuencias de este apartado.

2.2.4. Consecuencias

1. La vulnerabilidad natural del cuerpo humano representó el origen de la implementación de estrategias de supervivencia, tal como la cooperación.
2. La historia natural explica la presencia de instintos sociales (incluida la empatía como piedra angular), así como también, da cuenta *parcialmente* de los elementos que favorecieron la selección de mecanismos encargados de nuestras orientaciones normativas generales, lo que, en conjunto, constituye un núcleo biológico de rasgos heredables, *i.e.* la herencia genética.
3. El origen cognitivo del juicio moral se corresponde con la evolución de capacidades socio-cognitivas tales como *leer* (inferir) las intenciones de otros, por lo que, se ajusta bien con las capacidades atribucionistas de la cognición social.

Tras haber visto estas aproximaciones evolutivas, ahora, corresponde abordar los supuestos cognitivos que habilitan el marco conceptual computacionalista, en donde se verá claramente en qué consiste el procesamiento cognitivo-funcionalista.

2.3. Cognición: supuestos cognitivos computacionalistas

Primero, recordemos que la visión funcionalista-computacionalista basa su explicación cognitiva en un sistema intracraneal, representacional y regido por un conjunto de reglas o principios para procesar la información proveniente del entorno.

De ello, se ha considerado tradicionalmente que el modelo cognitivo que habilita el marco computacionalista es el llamado “modelo sándwich” (Susan Hurley, 1998). El modelo sándwich, en conformidad con la visión funcionalista, concibe la cognición como procesamientos de información que son regidos por reglas o principios y mediados por representaciones mentales. El modelo funciona de manera lineal:

input perceptual → cognición (procesamiento de información) → *output* motor/toma de decisiones

Bajo este modelo, el sujeto cognitivo se enfrenta a un mundo pre-dado que después representa, es decir, el sujeto percibe el mundo y tales percepciones (*inputs*) entran a los mecanismos de procesamientos de información, en donde son procesados por algoritmos bien definidos, lo que resulta en toma de decisiones que se convierten en acciones materializadas (la conducta). Este procesamiento es lineal y jerárquico, y básicamente funciona de la siguiente manera:

“Este enfoque, [...] enfatiza un marco de procesamiento jerárquico pero en serie, donde la información siempre fluye en una sola dirección: la información sensorial viaja hacia arriba en la jerarquía, la cognición ocurre en ese procesamiento de información y luego la información motora viaja hacia abajo en la jerarquía en forma de comandos musculares” (Carpendale, Frayn, y Kucharczyk, en Kiverstein, 2017, p. 198).

Ahora bien, de manera relevante, este modelo resulta en una separación modal de la percepción y la acción. El computacionalismo se centra en la manera en que el sistema cognitivo representa, procesa y transforma la información. Así, la "cognición" o "los procesos cognitivos" han sido entendidos como uno de los factores o causas endógenas que contribuyen a la producción del comportamiento (Aizawa, 2014). En contraste, los teóricos corporizados, en su mayoría, homologan la cognición con la conducta, por lo que no hay tal separación debido a la naturaleza de sus principios.

Así, pues, desde estas explicaciones funcionalistas, se ha asumido que la intervención del cuerpo en los procesos cognitivos es trivial y en realidad la mente superviene de procesos neuronales o cerebrales, resultando así en un enfoque neuro-centrista.

En suma, el computacionalismo entiende la cognición como cálculos o *computaciones* representacionales que implican procesamientos de información regidos por principios enfocados en la resolución de diversos problemas. Así, sus compromisos ontológicos y con diversas entidades teóricas son manifiestos: la cognición implica procesos informacionales algorítmicos sobre representaciones mentales.

Entonces, desde esta perspectiva, el procesamiento cognitivo-funcionalista funciona de la siguiente manera:

2.3.1. Procesamiento cognitivo- funcionalista:

input (perceptual) → cognición (procesamiento de información) → *output* (toma de decisiones que deriva en conductas manifiestas).

Así, todo fenómeno cognitivo que siga este modelo, será un procesamiento cognitivo-funcionalista.

Ahora bien, notablemente se verá que en la arquitectura *súper-modular* los procesamientos ya no son lineales, sino, multidireccionales y altamente interactivos. No obstante, dada su naturaleza computacional, preserva la separación modal entre la percepción y la acción.

Para mis propósitos, lo relevante del procesamiento cognitivo-funcionalista es notar que por su naturaleza funciona mediante entradas perceptuales (*inputs*) y procesamientos cognitivos (procesos de información) representacionales (lineales o multidireccionales) que generan toma de decisiones y conductas (*outputs*), manteniendo así dicha separación modal percepción-acción (mediada por la representación).

De ello, mi propuesta respecto a la cognición moral es que los juicios morales y las acciones morales que se expliquen mediante procesamientos cognitivos representacionales guiados por principios, pertenecerán al marco computacionalista, ejemplo: los juicios y las acciones morales que empleen estrategias socio-cognitivas que impliquen ToM y, por tanto, la perspectiva epistemológica de 3ª persona, así como también, el uso estricto de principios morales.

Bajo este supuesto, a continuación, analizo la propuesta de la cognición social computacionalista en su versión modularista-innatista, dado que a partir de esta versión es posible develar las limitaciones que deben ser superadas, para después explicitar la arquitectura *súper-modular* que es capaz de hacerlo por su característica altamente interactiva.

2.4. Cognición social computacionalista

Una visión computacionalista innatista se suscribe plausiblemente a los supuestos de la escuela de Santa Bárbara, lo que implica ser en gran parte *adaptacionista*, *modularista* y de cierta manera *aculturalista*. Entonces, aquí corresponde presentar los argumentos de tal visión, la cual es la cognición social modularista-innatista.

Los debates conceptuales pertinentes a la teoría de la mente se vieron con detalle en el capítulo 1, por lo que aquí me centro en presentar la versión mencionada.

Primero, recordemos que los teóricos nativistas plantean que en el cerebro existen módulos especializados en computar la información de cada una de las distintas capacidades

cognitivas, tales módulos son innatos y se desarrollan en función de la maduración biológica, de esta manera, existe un módulo especializado para la teoría de la mente (ToM) y éste puede componerse de más sub-módulos. La idea es que tenemos una capacidad modular convertida en una facultad para *computar* mentes, es decir, un mecanismo de especial atención que emplea conceptos innatos tales como *creencias* y *deseos*, conceptos innatamente específicos que predisponen el desarrollo normal de los infantes para prestar atención selectiva a los estados mentales, lo que explicaría cómo el infante aprende acerca de los estados mentales y cómo atribuirlos (Fodor, 1983, 1992; Leslie, 1987, 1994; Baron-Cohen, 1995).

Segundo, es conveniente aludir brevemente a los supuestos que están detrás de la visión computacionalista de la cognición social, la llamada cognición social atribucionista:

- 1) Explica la competencia socio-cognitiva como atribución mental.
- 2) Toma la perspectiva epistemológica de 3ª persona (en los encuentros sociales los individuos son *observadores*).
- 3) En consecuencia, el individuo observa, atribuye estados mentales y después interactúa.

Como se vio en el capítulo 1, las dos teorías reinantes de la atribución mental son la teoría-teoría (TT) y la teoría simulacionista (TS), las cuales, en cualquiera de sus versiones, siguen “el proceso de la cognición social computacionalista”.

- **Proceso de la cognición social computacionalista:**

El sujeto adscriptor *observa* al otro y percibe la información relevante (señales del entorno y conducta manifiesta) → procesa tal información mediante el proceso de atribución mental postulado (ya sea TT o TS, o una hibridación de ambos), lo que le permite producir una representación del otro y de sus estados intencionales → obtiene un estado resultante, el cual atribuye al otro. Con ello, se explica y predice la conducta del otro. Respecto a la acción social, es la representación del otro y de sus estados intencionales lo que le permite al sujeto adscriptor planificar cómo actuar de una manera socialmente apropiada, ante lo cual, responde ejecutando ese plan con comandos motores lo mejor posible.

Tras esta breve exposición de los supuestos generales de las teorías atribucionistas, ahora corresponde analizar la ToM modularista que asume, en particular, que el agente es un

observador que atribuye estados mentales en virtud de un módulo innato encargado de las tareas relevantes pertinentes a ToM.

2.4.1. Cognición social modularista

Como punto de partida, es necesario precisar ¿qué es un módulo? o bien ¿qué es ser modularista?

Existen diversas concepciones de “módulos” que resultan en distintas explicaciones. Una primera explicación detallada de “modularidad” es la que ofrece Jerry Fodor (1983), quien concibe los módulos como portadores de algunas, o de todas, las siguientes características: 1) son de dominio específico (i.e., tienen la función de resolver problemas en dominios particulares), 2) son de operación obligatoria (i.e., cuando una entrada (*input*) es recibida, ésta se procesa automáticamente de una forma particular e independiente al control voluntario; 3) su acceso central se limita a las representaciones mentales que *computan* los sistemas de entrada (i.e., el flujo de información *fuera* de un mecanismo está restringido); 4) poseen alta velocidad de cálculos o *computaciones*; 5) son encapsulados, o bien, su información es encapsulada (i.e., el flujo de información *dentro* de un mecanismo es restringido y no se traslapa con la información de otros módulos); 6) poseen una arquitectura neuronal fija; 7) tienen patrones de avería específicos, y; 8) poseen un ritmo específico en la ontogenia, tanto temporal como de secuencia.

Fodor postuló así la modularidad para explicar la cognición de *bajo-nivel*³³; en particular, los procesos que subyacen a las capacidades perceptuales y al procesamiento sintáctico del lenguaje (Bohl, 2014, p.46). Y, la cognición de *alto-nivel*, tal como la formación de creencias, la explicaba mediante procesos de propósito-general. No obstante, otros teóricos modularistas han ampliado esta concepción fodoriana y la han llevado a abordar la cognición de alto-nivel también, alejándose de la noción de “encapsulamiento informacional” (p. ej. Carruthers, 2006) y, en esta línea, incluso, hay autores que han separado la noción de “módulos especializados” de la noción de innatismo (p. ej. Karmiloff-Smith, 1992, 1998).

³³ Fodor defendía la existencia de módulos responsables de tareas simples, i.e. de bajo-nivel, como resultado de la evolución filogenética. Sin embargo, rechazaba la idea de que existieran módulos para tareas más complejas, i.e. módulos encargados de cogniciones de alto-nivel, pues consideraba que éstas se efectúan mediante procesos que los módulos simplemente no llevan a cabo por definición, procesos que no podrían ser de dominio específico, p.ej. el pensamiento y el razonamiento abductivo (Enesco y Delval, 2006, p.9).

Así, debido a los diversos debates y a las distintas concepciones, ahora el término “módulo” o “modularidad” se ha vuelto vago y brumoso, por lo que es necesario determinar de qué noción parte cada explicación.

La cognición social funcionalista tradicional toma como punto de partida la explicación de “la modularidad computacional”, la cual caracteriza un sistema computacional que convierte las representaciones de entrada (*input*) en representaciones de salida (*output*) mediante el uso de un conjunto de reglas sintácticas (ver Segal, 1996, pp. 143–145). Los módulos que le constituyen son en su mayoría estáticos y se corresponden con los principios fodorianos expuestos anteriormente³⁴.

La principal característica de los modularistas de la cognición social es que se basan en los sistemas innatos de procesamiento de información para explicar el desarrollo cognitivo y, en ese sentido, consideran que la selección natural favoreció un mecanismo o un módulo específico para las capacidades de teoría de la mente.

En particular, la versión innatista de la cognición social surge para resolver la cuestión de universalidad de la Teoría de la Mente que desarrollan los infantes, es decir, ¿por qué parece que la trayectoria de desarrollo de ToM es similar y universal entre los infantes neurotípicos? Para resolver esta cuestión, los innatistas argumentan que “la selección natural produjo mecanismos específicos que impulsan el desarrollo de teoría de la mente” (Bohl, 2014, p. 47). Con esta idea intentan asegurar la universalidad de ToM.

Sus principales exponentes son Alan Leslie y Baron-Cohen, respectivamente. En general, las teorías de ambos postulan diversos sub-módulos constituyentes de un módulo de teoría de la mente (ToMM) para explicar la competencia humana de lectura-de-mentes.

En breve, Leslie (1994) postula el desarrollo ontogenético de teoría de la mente como la maduración de un sistema de dominio-específico, es decir, del módulo de teoría de la mente (ToMM), que es la sede de las habilidades de la lectura mental o de la atribución mental. A su vez, propone tres sub-módulos que constituyen a ToMM: 1) *ToBy*, o teoría del

³⁴ No obstante, se verá que tal explicación ha evolucionado, pues ahora los planteamientos nativistas de la cognición social hablan de módulos que no sólo se desarrollan en función de la maduración biológica, sino que también aceptan procesos de cambio en el desarrollo.

Mecanismo Corporal; 2) *ToMM1*, o sistema 1 del Módulo de la Teoría de la Mente, y; 3) *ToMM2*, o sistema 2 del Módulo de la Teoría de la Mente.

Por su parte, Simon Baron-Cohen (1997) concede en general que ToMM habilita la atribución de la gama completa de los estados mentales. No obstante, explica los tres sub-módulos que le constituyen de manera distinta, además de que éstos se limitan a procesar estados perceptuales y volitivos: 1) detector de intencionalidad (ID); 2) detector de dirección de miradas (EDD), y; 3) mecanismo de atención compartida (SAM). Estos sub-módulos en interacción con ToMM generan la atribución mental correcta, según Baron-Cohen.

El postulado fundamental de las explicaciones innatistas-modularistas es que nuestras capacidades de atribución mental se dan en función de la maduración de mecanismos cognitivos innatos de dominio-específico, con lo que rechazan los postulados empiristas que sostienen que el desarrollo de las capacidades de atribución mental se da a través de una capacidad de dominio-general de construcción de teorías (como la “teoría del niño científico” que se vio en el capítulo 1 en la sección de “Antecedentes de la cognición social”).

En general, se les ha acusado a los innatistas de ser “teóricos del anti-desarrollo” o, en otras palabras, de ser incapaces de explicar, por ejemplo, el desarrollo de una secuencia típica de cambios conceptuales que caracteriza el desarrollo de una teoría de la mente competente, es decir, no son capaces de explicar cómo ocurre el reemplazo de una comprensión temprana no-representacional de los estados mentales por una explicación representacional posterior (Gopnik y Wellman, 1992, p. 284).

En contraste, los empiristas explican la transición del desarrollo conceptual por ofrecer una explicación dinámica que toma en cuenta la relevancia que poseen las interacciones entre el infante y el entorno, tanto para el cambio conceptual como para la construcción de teorías: las teorías cambian en un orden sistemático bajo la presión de nueva evidencia (Bohl, 2014, p.51).

Algunos innatistas, como Leslie, sortean esta crítica por conceder que los módulos no son estáticos respecto al desarrollo y aceptan que, si bien, existen restricciones para el flujo de información dentro y/o fuera de un módulo, esto no significa que los módulos no puedan *desarrollarse* internamente (Scholl y Leslie, 1999). De esta manera, la teoría de Leslie defiende la uniformidad de trayectorias del desarrollo de ToM, argumentando que las

diferencias pueden estar en el desempeño, pero no en la competencia. Entonces, ¿qué tipo de desarrollo estaría aceptando?

Aquí es crucial notar que introducir el aspecto de *desarrollo* en la noción de módulos y en particular en el módulo de teoría de la mente (ToMM), corre el riesgo de conducir a aceptar distintas trayectorias en su desarrollo, es decir, su desarrollo podría resultar en distintas teorías de la mente afectadas según el entorno, el aprendizaje, etc., lo que contraviene el supuesto universalista de los innatistas. Por tanto, se debe hacer explícito a qué tipo de *desarrollo* se refieren.

Para hacer frente a este desafío, el postulado innatista señala que hay módulos susceptibles de desarrollo, pero que tal desarrollo estaría determinado genéticamente; planteamiento con el que intentan asegurar la universalidad respecto a la trayectoria del desarrollo de ToM.

Pero, ¿qué ocurre entonces con el supuesto de que los módulos son estáticos? La diferencia clave que debe ser reconocida aquí es la concepción de “módulos susceptibles de desarrollo” y la concepción de “módulos estáticos”. Al respecto, Segal (1996) distingue dos dimensiones de modularidad: la modularidad *sincrónica* y la modularidad *diacrónica*. La primera caracteriza los módulos como mecanismos estáticos, en tanto que poseen una arquitectura neuronal fija, tienen patrones de avería específicos y poseen un ritmo específico en la ontogenia, tanto temporal como de secuencia (p.ej. la modularidad fodoriana), mientras que la segunda concibe los módulos como susceptibles de desarrollo.

Explico, la diferencia es que la modularidad *sincrónica* concibe que el desarrollo y la funcionalidad de los módulos dependen de un programa madurativo de base innata, que florecerá según un ritmo y una serie de estadios independientemente de la *experiencia* con el entorno. Así, los módulos se desarrollarían invariablemente según una secuencia característica (ontogenia universal), lo que supone incluso formas características de “desviación” o “ruptura” (patología universal) (ver Gutiérrez-Martínez, 2005, p.182).

En contraste, una explicación que aborda el tema del desarrollo de módulos es la explicación que Segal ha llamado “parametrización de módulos diacrónicos”; de ahí su nombre de “modularidad *diacrónica*”. En particular, los módulos diacrónicos son módulos que se desarrollan con el tiempo siguiendo trayectorias genéticamente determinadas, pero que aceptan ciertas intervenciones de datos perceptuales del entorno. Según Segal, un

ejemplo de ello es la explicación de Chomsky concerniente al desarrollo del lenguaje; al respecto, dice Segal: “El aprendizaje del lenguaje en la infancia se puede explicar como un proceso de *parametrización* (establecimiento de parámetros) dentro de límites genéticamente predeterminados: de acuerdo con la información recibida (*input*) del entorno lingüístico, los niños adoptan implícitamente las reglas de ese lenguaje en particular” (Segal, 1996, pág. 146). Así, los niños en diferentes entornos lingüísticos terminan hablando diferentes lenguajes. De acuerdo con Segal, esto puede explicarse como resultado de un proceso diacrónico de ajustar los parámetros provenientes del entorno lingüístico con los límites genéticamente especificados del módulo lingüístico (Bohl, 2014, p.51); supuesto que asegura la universalidad de dicha capacidad.

Pero, ante estos planteamientos, cabe cuestionar ¿qué papel juega exactamente el entorno o el aprendizaje en el modularismo?

Para ser clara, una versión fuerte del modularismo concibe que los módulos están preadaptados y no sufren, en sentido estricto, ningún desarrollo, sino sólo un “enriquecimiento” y, en este contexto, se considera que el aprendizaje y el entorno son meros factores *desencadenantes* de las funciones que vienen previamente determinadas, los cuales pueden adelantar o retrasar el desarrollo, pero no alterar su curso ni su forma final. Desde esta perspectiva, entonces, “el aprendizaje consiste en un enriquecimiento de los principios básicos y su consolidación [...]” (Carey y Spelke, 1994, p. 243, citado en: Enesco y Delval, 2006, p.251), no habiendo así alteraciones ni en el curso de desarrollo ni en la forma final de la capacidad en cuestión. Con ello, se pretende asegurar la universalidad de las trayectorias de desarrollo, ya que las fallas y diferencias estarían en el desempeño, pero no en la competencia en sí.

En paralelo, Chomsky ha sostenido que la capacidad lingüística está genéticamente programada en los seres humanos y que no existe un proceso de *aprendizaje* del lenguaje propiamente dicho, sino un proceso de “maduración biológica”, similar al proceso de *aprender* a caminar o de *aprender* a volar, en el caso de las aves (Ramos-Arenas, 1992, p.28). Según Chomsky, existe una suerte de “órgano del lenguaje” en cuyo desarrollo se funda nuestra adquisición de la capacidad lingüística, en donde el desarrollo lingüístico está determinado genéticamente y consiste en un proceso de “maduración” más que de “aprendizaje”. Una implicación de ello es notar que, bajo esta concepción, dadas las condiciones necesarias, el niño prácticamente no puede evitar aprender su lengua (Ramos-Arenas, 1992, pp. 27-28).

Ahora bien, una interpretación útil que hace Fodor acerca de la noción chomskyana de “aprendizaje” es la siguiente:

“Cuando él dice que la teoría del aprendizaje del lenguaje es la historia de cómo madura la facultad del lenguaje, lo que quiere decir primariamente es que la ontogenia de las capacidades lingüísticas es el despliegue de las consecuencias deductivas de creencias innatas en interacción con un cuerpo de datos perceptuales” (Fodor, 1983, p. 7).

Así, pues, bajo este enfoque, el aprendizaje se homologa con el desarrollo y se explica en términos meramente madurativos y, por su parte, la experiencia sensible o el contacto con las cosas del mundo son factores *desencadenantes* de lo que yace innatamente.

Ciertamente, la diferencia entre las dimensiones de modularidad *sincrónica* y de modularidad *diacrónica* radica en el papel que le otorgan a la “experiencia con el entorno” en sus explicaciones. Mientras que la modularidad *sincrónica* sostiene que el desarrollo y la funcionalidad de los módulos florece según un ritmo y una serie de estadios independientemente de la “experiencia con el entorno”, la modularidad *diacrónica* la reconoce y le da un papel de *desencadenante*. No obstante, ambas conciben el desarrollo como procesos de maduración biológica.

En general, los planteamientos modularistas expuestos presentan tensiones que usualmente derivan en las siguientes objeciones (ver Gutiérrez-Martínez, 2005, p.183):

- 1) El modularismo concibe el desarrollo en términos puramente madurativos, lo que implica que las diferencias que se contemplan entre el niño y el adulto respecto a la facultad en cuestión, sólo sean de carácter *cuantitativo*, es decir, el infante posee los mismos dispositivos que el adulto, pero ellos aún no son completamente funcionales.
- 2) Debido a la especialización de los módulos, las posibilidades del conocimiento (conceptos, creencias, etc.) que podemos desarrollar son limitadas y están previamente condicionadas.
- 3) El aprendizaje está ligado a dotaciones prefijadas, lo que implica que la experiencia tenga un papel mínimo de mero *desencadenante* y el aprendizaje esté seriamente constreñido.

Ahora bien, una objeción específica a la cognición social modularista la presentan Currie y Sterelny (2000, p. 149) quienes argumentan que “es poco probable que la lectura de mentes sea fuertemente modular (en el sentido de ser el producto de mecanismos cognitivos

informacionalmente encapsulados) porque, para formular creencias sobre los estados mentales de otras personas, la información contextual específica debe ser (y a menudo es) tomada en cuenta” (en Bohl, 2014, p.52).

Entonces, las objeciones se enfocan principalmente en cuestionar la especialización y el encapsulamiento informacional de los módulos, además de la concepción de aprendizaje tan poco flexible que resulta de sus postulados, pues, si asumimos el aprendizaje y la experiencia como meros *desencadenantes* de funciones previamente determinadas ¿cómo explicaríamos el cambio y la enorme diversidad adaptativa del ser humano a lo largo de su historia, en diferentes culturas y contextos? (ver Enesco y Delval, 2006). Y en esta línea, ¿cómo podríamos hacer compatible al modularismo con la evidencia irrefutable de la plasticidad cerebral? pues si fuera el caso de una modularidad así definida, habría que sumar a nuestro aparato cognitivo un sinnúmero de mecanismos para dar lugar a estrategias adaptativas capaces de hallar el mejor curso de acción o la mejor decisión a problemas específicos en contextos particulares, lo que lo haría computacionalmente ineficiente además de que no se podría explicar lo bien que los humanos aprenden a resolver nuevos retos medioambientales (ver Hernández-Chávez y García-Campos, 2020, p.62). En otras palabras, no se podrían explicar los reajustes que el entorno exige en los diversos contextos.

Así, pues, estos planteamientos imponen al menos dos claros desafíos al modularismo: 1) necesita involucrar información pragmática (información del contexto) de manera mucho más sustancial y 2) requiere soportar fuertes interacciones entre los diversos mecanismos cognitivos para efectuar con éxito las tareas relevantes.

Para hacer frente a estos desafíos, requerimos una arquitectura cognitiva capaz de superar estas limitaciones, al mismo tiempo que cuente con apoyo empírico consistente.

Al respecto, en la actualidad contamos con una propuesta sólida capaz de satisfacer estas demandas y por tanto presenta ventajas teóricas y metodológicas frente a otros planteamientos disponibles, a saber, la arquitectura *súper-modular* (García-Ramírez, 2019). La arquitectura súper-modular es una arquitectura cognitiva representacional, potencialmente interactiva, multidireccional y que involucra, de manera relevante, información pragmática y razonamiento heurístico en sus procesamientos, por lo que es capaz de explicar satisfactoriamente los procesamientos computacionalistas implicados en la cognición moral (dicha propuesta la presento más adelante).

Ahora bien, una vez efectuado el análisis de las tres secciones abordadas hasta aquí, en donde presenté los supuestos computacionalistas pertinentes a la evolución, a la cognición y a la cognición social, ahora corresponde abordar concretamente la cognición moral computacionalista.

Previo a ello, permítaseme hacer un recuento de las consecuencias relevantes de lo que se ha visto en este capítulo, esto con el fin de evitar inconsistencias.

2.4.2. Consecuencias

1. Respecto a la evolución, se vio que la selección natural contribuye a explicar la presencia de un núcleo biológico de rasgos heredables que incluiría instintos sociales, empatía y mecanismos seleccionados que subyacen a las orientaciones normativas generales exitosas y/o perniciosas, pero ello no basta para dar cuenta del espectro completo de la moral, por lo que falta explicar las intervenciones de la evolución cultural en la evolución de la misma.
2. Respecto a la arquitectura cognitiva que propongo, postulo los dos modos de cognición: el computacionalista y el enactivo (cada uno con su metodología y su entidad teórica núcleo). En relación con el computacionalismo, destacué los problemas que presenta el modularismo en general, por lo que planteé que es necesario recurrir a una arquitectura súper-modular capaz de superarlos; arquitectura cognitiva que debe ser especificada. En esencia, es un sistema representacional *abierto*, multidireccional y altamente interactivo.
3. Acotación: se verá que una consecuencia de asumir la arquitectura súper-modular es que el procesamiento cognitivo-funcionalista ya no es lineal, de abajo a arriba y ejecutado hacia-adelante, sino que ahora es multidireccional y se beneficia de múltiples recursos. No obstante, tómesese en cuenta que por su naturaleza computacionalista, preserva la separación modal entre la percepción y la acción.

Con estas cuestiones en mente, doy paso a la propuesta computacionalista de la cognición moral.

2.5. Cognición moral computacionalista

En esta sección abordo dos temas centrales, el cognitivo y el normativo.

Respecto a lo cognitivo: asumo la arquitectura súper-modular y analizo los procesamientos que subyacen a nuestras facultades morales, por lo que gran parte de esta sección se

centra en exponer la evidencia empírica proveniente de la psicología del desarrollo moral y, en específico, en explicitar los hallazgos encontrados en infantes pre-lingüísticos para apuntalar el argumento del conocimiento moral general dotado evolutivamente.

Respecto a lo normativo: mi propuesta es que los humanos compartimos un tipo de conocimiento moral general capaz de orientarnos normativamente hacia nociones universales de lo correcto/incorrecto; lo que aseguraría su objetividad y explicaría por qué toda población humana tiende a desarrollar, al menos, nociones de un sistema moral.

Así, la estructura de esta sección es la siguiente: primero, presento una suerte de “Justificaciones” en las que clarifico las diferencias puntuales entre la propuesta del núcleo moral innato (NMI) y mi propuesta, así como también, especifico de qué tipo de *conocimiento* moral general estoy hablando. Segundo, reviso la arquitectura súper-modular, en donde se verá que es computacionalista y altamente interactiva en el sentido de que apela a una interacción multi-modular que es multi-direccional y soporta la interacción entre los aspectos de dominio general y de dominio específico, incluso, integra sustancialmente la información pragmática y el razonamiento heurístico para efectuar la tarea relevante (tales conceptos serán explicados más adelante), lo que hace justicia a lo indeterminado de cada contexto. Visto esto, en tercer lugar, expongo los problemas heredados del modularismo tradicional y trazo brevemente cómo la arquitectura súper-modular es capaz de abordarlos, lo que posibilita ver claramente su potencialidad. En cuarto lugar, me centro en el tema de las facultades morales y presento el amplio cuerpo de trabajo empírico que apuntala los planteamientos innatistas, debido a que su evidencia permite argumentar a favor del conocimiento moral general que postulo. Por último, expongo las implicaciones normativas que derivan de los supuestos computacionalistas aquí asumidos, según mis consideraciones.

2.5.1. Justificaciones

- Diferencia entre NMI y mi propuesta

Es preciso recordar que lo que habilita mi propuesta de un marco integrativo es conceder que hay dos tipos de modos cognitivos (computacionalista y enactivo), los cuales implican dos tipos distintos de metodologías (individual/interna y relacional) y dos entidades teóricas núcleo que habilitan a cada una de ellas (representación y acción: *representación* para el computacionalismo y *acción* para el enactivismo).

En esta línea, por un lado, defiendo que el computacionalismo se enfoca en explicar los procesamientos cognitivos representacionales de los individuos, que son en gran parte sub-personales y, por el otro, argumento que es necesario contemplar el enactivismo para explicar el nivel supra-personal como un dominio autónomo desde su aspecto *relacional*.

Así, la diferencia puntual entre el NMI y mi propuesta es que el NMI explica el nivel supra-personal en términos de procesamientos cognitivos representacionales. En contraste, en mi propuesta le otorgo autonomía y lo defino desde una metodología relacional, de la que emergen estructuras que constituyen significados y normas que hacen de las interacciones un lugar donde la vida social puede ocurrir; y por lo tanto dan lugar a capacidades cognitivas completamente nuevas (Steiner y Stewart, 2009, p. 536).

Por consiguiente, mi propuesta es que contamos con una base universal representacional de la moral común entre los seres humanos, que se compone de principios morales pre-lingüísticos y es fuertemente influenciada por otros mecanismos y dominios cognitivos que involucran información pragmática de manera relevante; pero también, defiendo que hay una dimensión moral que emerge de una epistemología relacional, pues es esto lo que da lugar a *algunos* significados y normatividades *compartidas* que surgen de la co-constitución de las interacciones y de las prácticas sociales irreductibles a los procesos individuales del sujeto cognitivo.

Se trata, pues, de estudiar la moral en dos niveles: 1) desde los procesamientos cognitivos representacionales que le constituyen, y, 2) desde los procesamientos relacionales que la develan como una estructura social instituida en la práctica social.

En consecuencia, mi propuesta integrativa tiene dos propósitos: explicar (1) mediante la *representación* (entendida al menos como una estructura portadora de información) y explicar (2) mediante la *acción*, entendida ésta como acoplamientos dinámicos agente-entorno o agente-agente que generan conductas que implican responsabilidades (Steiner y Stewart, 2009, p.529).

- **¿Qué justifica que requerimos ambos niveles?**

Recientemente, en la literatura de la cognición social, algunos autores han proporcionado diferentes argumentos a favor de fusionar elementos “atribucionistas” y elementos “interaccionistas” en lugar de preferir una cuenta sobre la otra (Bohl y van den Bos, 2012; Gangopadhyay y Schilbach, 2011; Michael, 2011; Bohl, 2014). Concretamente, los autores

defienden que ambos elementos y metodologías explican *distintos* aspectos de la cognición social, por lo que su fusión da lugar a una visión más abarcadora de la misma: la atribución mental explica la parte socio-cognitiva que es particularmente humana, es decir, la capacidad del razonamiento abstracto sobre los estados mentales de otras personas, mientras que los interaccionistas dan cuenta de los elementos perceptuales y de los procesamientos supra-personales que trascienden a los individuos (Bohl y van den Bos, 2012, p.4)³⁵.

De acuerdo con los autores, ambas metodologías contribuyen significativamente a la explicación de la cognición social. Según los interaccionistas, la metodología computacionalista requiere de una metodología relacional ya que la interacción social no podría explicarse enteramente por referir a los procesos representacionales que se desarrollan en la mente (o el cerebro) de los individuos, sino que para comprenderla, habría que considerar también el proceso de la interacción en el nivel supraindividual como un genuino dominio cognitivo que se desarrolla *entre* sujetos (Reddy y Morris, 2004, p. 653) y de esta manera contemplar los significados co-creados que de ello surgen. Asimismo, sostienen que el nivel interactivo debe estimarse como algo que no es reducible al *input* del proceso cognitivo de cada individuo (De Jaegher et al., 2010), (ver Bohl, 2012, p. 3). En otras palabras, se trata de revelar el aspecto *relacional* de la cognición social.

En sentido concomitante, los recientes desarrollos empíricos (neurocientíficos) de la competencia socio-cognitiva también han señalado la necesidad de sumar a las teorías disponibles (tales como las computacionalistas) investigaciones que estudien los encuentros sociales desde una perspectiva de segunda persona, en tiempo real y de una manera verdaderamente *interactiva* (Schilbach, Timmermans, Reddy, Costall, Bente, Schlicht, y Vogeley, 2013), es decir, de una manera *relacional*.

Asimismo, la metodología interaccionista, por su parte, requeriría de una metodología interna porque es necesario conocer los mecanismos individuales (Bohl y van den Bos, 2012, p.4), tales como los procesamientos de información representacionales guiados por principios que dan lugar al fenómeno cognitivo de interés. Para hablar de la cognición socio-moral en toda regla se requiere conocer los procesamientos y mecanismos cognitivos

³⁵ La arquitectura cognitiva que proponen los autores está basada en un sistema dual que apela a una explicación de naturaleza mecanicista del tipo Bechtel (2008), capaz de soportar múltiples capas para diferentes niveles de explicación. En el capítulo 4 hago explícita la diferencia entre esta propuesta y la arquitectura cognitiva que propongo.

representacionales que la habilitan y dan lugar a ella, los cuales son capaces, entre otras cosas, de abordar el razonamiento abstracto sobre los estados mentales de otras personas, pues el aspecto relacional del nivel supraindividual no es suficiente para explicarlo. En otras palabras, desde esta perspectiva, la metodología interaccionista se complementaría bien con la computacionalista debido a que la cognición socio-moral también trata de procesamientos de información guiados por principios.

En adición, según los autores, otro aspecto vulnerable del interaccionismo es que tiende a enfatizar demasiado las interacciones sociales que son honestas, fluidas y cooperativas. Cuando es el caso de que la interacción social evoluciona de manera agradable para todos los participantes, seguramente hay poca necesidad de que los interactuantes atribuyan estados mentales (a los otros) basados en simulación o inferencia. Pero es cuestionable si éste es el caso en situaciones de competencia, desacuerdo, conflicto o malentendidos evidentes (Bohl y van den Bos, 2012, p.5); tales como en situaciones problemáticas y conflictivas desarrolladas en contextos moralmente relevantes. Esto requerirá más que la percepción directa de los estados de otros, pues “es difícil ampliar la característica de perceptibilidad a *todos* los estados mentales, ya que es obvio que algunos estados mentales (p. ej., pensamientos que no se expresan) son menos perceptibles que otros (p. ej., emociones fuertes), incluso según los estándares interaccionistas” (*Íbid*).

Por otra parte, Varela (1999) y otros teóricos han reconocido la distinción entre los juicios y las justificaciones morales por un lado y el comportamiento moral (inmediato) por el otro, siendo el segundo tema el que menos atención ha recibido, por lo que parece necesario ofrecer explicaciones del mismo, a lo que puede contribuir significativamente una perspectiva *relacional* que se base en el análisis de un confrontamiento inmediato con el mundo. A estos fines, dice Varela:

“Mi interés por el afrontamiento inmediato no significa que niegue la importancia de la deliberación y el análisis. Mi punto es que es importante comprender el papel y la relevancia de ambos modos cognitivos” (Varela, 1999, p.5).

Agregando que:

“Es esta inmediatez de percepción y acción la que queremos examinar críticamente. Este enfoque contrasta marcadamente con la forma habitual de investigar el comportamiento ético, que comienza analizando el contenido intencional de un acto y termina evaluando la racionalidad de juicios morales particulares” (*íbid*, p.2).

Se trata, pues, de abordar los dos modos de cognición y analizar las aportaciones que cada uno ofrece al tema de interés, dado que cada uno se enfoca en *distintos* aspectos: el

computacionalismo trata con el análisis intencional, representacional y con la acción guiada por principios y, el enactivismo, con los acoplamientos sensoriomotores dinámicos y con la acción guiada por la percepción que emerge del confrontamiento directo con el mundo, por tanto, como es de esperarse, ambos derivan en consecuencias distintas.

Concretamente, se verá que una consecuencia crucial es que el modo cognitivo guiado por principios explica la competencia moral (automática, sin esfuerzo y de uso simple) apelando principalmente al razonamiento *heurístico*, mientras que, el otro modo cognitivo guiado por la percepción, apela a lo que se conoce como *cognición ultrarrápida*.

En breve, la noción de *cognición ultrarrápida* “refiere a un comportamiento significativo que ocurre demasiado rápido para que el *procesamiento de información* en el cerebro desempeñe un papel en el control de cómo se desarrolla el comportamiento. El consenso es que el agente está preparado para ser activado por sus circunstancias de tal manera que automáticamente convergerá en el atractor conductual correcto.” (Mojica y Froese, 2019, p.3) De esta manera, “[l]a cognición ultrarrápida puede explicarse por sinergias que abarcan la mente y el cuerpo. Las sinergias son estructuras dinámicas temporales que anticipan un comportamiento apropiado al contexto. Un estado anticipatorio equilibra la mente y el cuerpo en simetría entre opciones de comportamiento equivalentes, y sólo un cambio mínimo en el contexto, favoreciendo una opción sobre cualquier otra, es suficiente para romper la simetría y activar una respuesta cognitiva ultrarrápida” (Wallot y van Orden, 2012, p.1).

Y por su parte, el razonamiento heurístico refiere al uso de “heurísticas rápidas y frugales” que son *reglas simples* en la caja de herramientas adaptativas de la mente para tomar decisiones con recursos mentales realistas (ver Todd y Gazzinger, 2000), las cuales favorecen la implementación de estrategias sencillas para la toma de decisiones rápidas y económicas en contextos limitados informacionalmente (más adelante se verá este concepto con más detalle).

Es conveniente ver los tiempos de reacción aproximados que se han demarcado respecto a cada una: los procesamientos de información como tal se ejecutan en no menos de *1 segundo* (Varela, 1999; Wallot y van Orden, 2012), las heurísticas de este tipo se discriminan en tiempos de reacción que van de *300ms* a *7000ms* (Ferreira, 2003), y la cognición ultrarrápida se efectúa desde los *150 ms*, o incluso menos (Wallot y van Orden, 2012).

En este punto las sutilezas se vuelven importantes, pues estas pequeñas diferencias derivan en distintas consecuencias explicativas respecto a los recursos que le habilitan. Se trata de abordar, por un lado, la intervención de *heurísticas rápidas y frugales* que se basan en el uso de reglas para el habilitamiento de estrategias sencillas en la toma de decisiones, las cuales forman parte del bagaje mental propio de la especie humana, y, por otro lado, de abordar la *cognición ultrarrápida*, la cual es guiada por la percepción y favorecida por la organización dinámica de sistemas sensoriomotores que ponen en marcha los esquemas sensoriomotores del agente.

Así, las distintas consecuencias radican en cómo conceptualiza cada marco conceptual el *saber-cómo (know-how)*³⁶ en la competencia moral, al menos en un mapeo de *grano grueso*: en el computacionalismo es automático, sin esfuerzo y basado en principios, y en el enactivismo, es inmediato, guiado perceptualmente y apela a la organización dinámica del sistema sensoriomotor (en donde se toman en cuenta las contribuciones del agente y del entorno en el desempeño cognitivo. Al respecto, Järvillehto (1998) sostiene que el estímulo, el contexto y el conocedor no hacen contribuciones separadas o causalmente disociables en el desempeño cognitivo³⁷).

Pero, ¿cuál es la diferencia entre ellos al momento de su aplicabilidad en situaciones moralmente relevantes? Dar respuesta a ello es una cuestión empírica y requerirá de mayor investigación³⁸, no obstante, aquí la idea es hacer notar el uso del razonamiento heurístico (guiado por principios) en la competencia moral y, por otra parte, la intervención de los acoplamientos sensoriomotores dinámicos en donde los esquemas sensoriomotores del agente son puestos en marcha, lo que da lugar al llamado “saber moral pre-reflexivo”

³⁶ Ryle (1946) distingue entre dos tipos de conocimiento: 1) un conocimiento práctico (*know how*) que se corresponde con un *saber cómo* hacer las cosas (ej. *saber cómo* andar en bicicleta), y; 2) un conocimiento proposicional o teórico (*know that*) que se corresponde con proposiciones que pueden ser verdaderas o falsas (ej. tener la creencia acerca de algo que es verdadero, digamos, *saber que* la fórmula química del agua es H₂O).

³⁷ “Timo Järvillehto (1998) ha esbozado una teoría del sistema organismo-ambiente que sitúa al contexto en un papel más fundamental. La cuestión crucial en la que se origina la teoría de Järvillehto sobre el sistema organismo-ambiente es si el organismo y el medio ambiente son dos sistemas causalmente distintos que de algún modo entran en contacto entre sí, o partes interdependientes de un único sistema integrado. La respuesta tiene implicaciones tanto conceptuales como empíricas porque Järvillehto sostiene que el estímulo, el contexto y el conocedor no hacen contribuciones separadas o causalmente disociables en el desempeño cognitivo” (Wallot y van Orden, 2012, p.7).

³⁸ Poe ejemplo, al momento, hay evidencia que muestra que los seres humanos *percibimos y clasificamos* las acciones como dañinas en milisegundos: el asesinato y la violación en 600 ms, el adulterio y el abuso en 700 ms, la prostitución y la mentira en 800 ms y finalmente la falta de respeto y los chismes en 850 ms. Estas acciones dañinas las *juzgamos* como inmorales en los mismos tiempos de reacción (Gray, 2018).

(Varela, 1999; De Souza, 2013). Esto lo destaco con la intención de revisar las intervenciones de cada modo-cognitivo en la cognición socio-moral.

Si bien, esto no representa una justificación conceptual completamente limpia, definiendo que al menos invita a explorar las ventajas que cada metodología puede ofrecer al resaltar sus elementos habilitadores, lo que a su vez busca otorgar un sitio a los dos modos-cognitivos.

Entonces, en mi propuesta, tanto los procesos cognitivos representacionales guiados por principios como el aspecto perceptual y relacional del nivel supra-personal son parte constitutiva de la cognición moral.

Una última aclaración es pertinente a especificar el tipo de “conocimiento” al que me refiero respecto al conocimiento moral general que propongo.

- ¿De qué tipo de “conocimiento” estoy hablando?

Se trata de hablar de un tipo de conocimiento psicológico sub-personal dotado parcialmente por la evolución, el cual funge como herramienta evaluativa objetiva capaz de orientarnos normativamente en contextos moralmente relevantes; lo que constituiría esa estructura sistemática moral compartida universalmente capaz de asegurar la objetividad moral.

La revisión de los argumentos evolutivos permite sugerir la presencia de predisposiciones morales dotadas en nuestra naturaleza cognitiva humana. Al respecto, se verá que la evidencia empírica lo apunala por sugerir la existencia de principios morales prelingüísticos (no aprendidos) y de predisposiciones naturales para comportamientos prosociales, entendimiento moral y evaluación moral.

En este contexto, definiendo que contamos con predisposiciones naturales para desarrollar facultades morales guiadas por aquellos principios morales prelingüísticos comunes entre los humanos, que con su desarrollo toman la forma de un “cuerpo de información” capaz de orientarnos normativamente³⁹, lo que constituiría los fundamentos morales universalmente compartidos. Los postulo como universales en tanto que tenemos la predisposición para desarrollarlos a partir de los mismos principios. Así, la maduración

³⁹ En la ciencia cognitiva tradicional, desde Fodor (1968), Nichols y Stich (1995), se ha ofrecido una noción de “conocimiento” entendido como un “cuerpo de información que guía alguna tarea”, el cual es implícito o sub-personal en el sentido de que el sujeto no es necesariamente consciente de poseerlo, por lo que, desde estas explicaciones, dicho conocimiento moral general plausiblemente puede pensarse como un cuerpo de información sub-personal que nos guía normativamente en contextos moralmente relevantes.

biológica tendría un papel fundamental, aunque incompleto, en el desarrollo moral (Hamlin, 2015, p.7).

Para dar cuenta de ello recurro a la explicación del NMI que concibe el conocimiento moral en términos de competencia moral, es decir, de acuerdo con el NMI el conocimiento moral se alcanza con el desarrollo de la competencia moral, por lo que, hablar de un conocimiento moral en toda regla requerirá del desarrollo de múltiples capacidades.

Desde este punto de vista, la competencia moral resulta de poseer y/o ejercer un conjunto complejo de habilidades que incluyen ToM, comprensión de intenciones, evaluación social y razonamiento bajo principios morales. En esta línea, se sostiene que “los sujetos conocen la moralidad en virtud de poseer dicha capacidad cognitiva basada-en-principios como parte de su maquinaria cognitiva moral” (García-Ramírez, 2019, en borrador, p. 273).

Así, desde este enfoque, el conocimiento moral es la capacidad de ejercer una cognición moral competente y la cognición moral resulta de la interacción entre múltiples mecanismos y procesamientos como los mencionados. Entonces, el conocimiento moral es una consecuencia del uso competente de esta capacidad cognitiva que es parte natural de nuestra dotación psicológica.

De manera relevante, el NMI postula que la posesión de esos principios morales otorga a los sujetos competentes la capacidad de aplicarlos en múltiples contextos y en un sinnúmero de escenarios. Por tanto, los sujetos competentes *saben cómo* participar en el pensamiento y la acción moral, puesto que el conocimiento moral así definido sería automático, rápido y de uso simple (ver García-Ramírez, 2019, en borrador, p. 273).

Por consiguiente, es un conocimiento del tipo saber-cómo, puesto que, concebir el conocimiento moral como competencia moral implica concebirlo como un saber-cómo. Sin embargo, también se corresponde con un *saber-qué* proposicional:

“La competencia moral incluye no solo las habilidades cognitivas, sino también los estados de información/representación constituidos principalmente por los principios de la cognición moral. Como tal, la competencia moral también sirve como fundamento para el conocimiento moral proposicional, ya que los principios de la cognición moral pueden fundamentar la verdad de una u otra proposición [...]. Así, aunque la competencia moral es una especie de saber-cómo, también hay un correspondiente saber-qué” (*Ibid.* p. 274).

Nótese que el conocimiento moral así definido es capaz de justificar parsimoniosamente la objetividad moral, debido a que, desde este punto de vista, los hechos morales son hechos psicológicos que son parte de nuestra maquinaria cognitiva, por lo que no tienen un origen *sui generis* ni se postulan como entidades externas a nuestra dotación natural, de tal manera que, sin este conocimiento moral no habría forma de determinar la estructura general de lo que es moral y por tanto no habría forma de ensamblar a partir de una base objetiva los códigos que propiamente pueden llamarse morales, al menos en mi propuesta.

Ahora bien, tras exponer estas justificaciones, ahora doy paso a revisar la arquitectura cognitiva súper-modular, para después presentar la evidencia empírica que apuntala las predisposiciones de los infantes pre-lingüísticos para las facultades morales.

2.5.2. Arquitectura súper-modular: superando las restricciones de los viejos debates

Primero, precisamos de una explicación cognitiva e interactiva para dar cuenta de la cognición moral, es decir, que vaya más allá de la concepción computacionalista acerca de que el sujeto es pasivo y sólo está a la espera de recibir los *inputs* del entorno para que se le *desencadene* el conocimiento de dominio-específico que yace innatamente.

Como mencioné, se trata de superar los viejos debates y hablar ahora de una capacidad cognitiva que se beneficia de la interacción entre diferentes mecanismos y dominios, específicos y generales. Tal capacidad ha sido apuntalada experimentalmente y es la llamada súper-modularidad (García-Ramírez, 2019). La explicación súper-modular es para hablar de una capacidad altamente interactiva que permite las influencias e intervenciones de diversos elementos como lo son distintas piezas de conocimiento, diferentes dominios cognitivos y distintos mecanismos para dar cuenta del fenómeno a estudiar.

Además, permite la *influencia* de otros elementos no atribuibles al algoritmo en función, tales como las heurísticas rápidas y frugales. Considero que esta concepción es innovadora y valiosa porque presenta ventajas ante las restricciones de las teorías tradicionales, debido a que, además de ser altamente interactiva, es multi-direccional e involucra información

pragmática⁴⁰ y estrategias heurísticas determinadas externamente para dar cuenta de lo indeterminado de cada contexto.

Entonces, el NMI se sitúa en una arquitectura súper-modular así definida, con lo que es capaz de superar las limitaciones y las restricciones de los viejos debates. Asimismo, desde esta perspectiva se plantea que la moral está sujeta a procesos de desarrollo de cambio y no solo de maduración (García-Ramírez, 2019, p.9).

De esta manera, el NMI parte de bases innatas, pero requiere contundentemente de las influencias del entorno para el desarrollo moral competente, en donde otorga un papel fundamental a las interacciones sociales y al uso de otras estrategias como lo es el razonamiento heurístico.

En este punto es conveniente instanciar cómo define el autor la *súper-modularidad* y su característica *interactiva*. Cabe mencionar que el ejemplo es respecto al lenguaje debido a que el autor se enfoca primeramente en explicar el “lenguaje natural”, ya que, en analogía con ello, explica la moralidad (más adelante se verá a este tema). No obstante, aquí es útil extraer su naturaleza característica:

Súpermodular: Para cualquier hablante S y cualquier lenguaje natural \mathcal{L} , S es un hablante competente de \mathcal{L} sí y solo sí S posee y/o ejerce un conjunto de múltiples y variadas habilidades modulares y de dominio general, específicas al lenguaje y no-lingüísticas, gracias a las cuales S adquiere, desarrolla y mantiene su conocimiento de \mathcal{L} .

Interactiva: para cualquier hablante S , cualquier lenguaje \mathcal{L} , con un lexicón λ , una sintaxis σ y una semántica Φ , S es un hablante competente de \mathcal{L} sí y solo sí los conocimientos de S acerca de λ , σ y Φ interactúan entre ellos, como también con información de dominios no lingüísticos tales como entendimiento de intenciones, toma de perspectiva, teoría de la mente y análisis estadístico entre otros.

(García-Ramírez, 2019, p.95)

⁴⁰ Refiere a la información del contexto que tiene que ver, entre otras cosas, con las limitaciones informacionales, prácticas y económicas, en donde se busca la interpretación más apropiada y económica de una situación.

Pero, respecto a la moral, ¿cuáles serían los procesamientos cognitivos involucrados? Se plantea que los constituyentes principales de la cognición moral son: Teoría de la mente, principios morales prelingüísticos, entendimiento intencional, entendimiento emocional y empatía. Veamos esto con más detalle.

2.5.3. Moral súper-modular

Primero, desde este punto de vista, el lenguaje se concibe tanto como un sistema representacional para propósitos comunicativos como también como un sistema que habilita la cognición humana, es decir, como una plataforma cognitiva que da pie a la cognición de orden superior. En este sentido el lenguaje (el cual es consecuencia y producto de una mezcla de distintos mecanismos) sería la plataforma potenciadora de la moralidad.

La propuesta del autor se basa en un análisis exhaustivo de un extenso cuerpo de evidencia empírica que muestra que el lenguaje no resulta de dominios específicos del lenguaje (lexicón, sintaxis y semántica) como se ha pensado tradicionalmente, sino que, en cambio, resulta de la interacción entre diversos mecanismos no específicos del lenguaje, tales como ToM, toma de perspectiva, razonamiento probabilístico, entendimiento de emociones, razonamiento práctico, conocimiento general y recursos cognitivos modulares, entre otros, además de que otorga un importante rol al uso de estrategias heurísticas rápidas y frugales para explicar la semántica. Al respecto, la semántica ya no la concibe como un asunto concerniente a las condiciones de satisfacción en correspondencia con el mundo (al menos no como su objetivo principal, pues ellas no son suficientes y requieren de elementos contextuales), sino como un procedimiento de toma-de-decisiones favorecido por razonamiento heurístico rápido y frugal, cuyo objetivo es encontrar la mejor interpretación que se ajuste a las limitaciones informacionales, prácticas y económicas de una situación, es decir, busca la interpretación más apropiada y económica de una oración en determinada situación (ver García-Ramírez, 2019, p.113).

Con esto, la arquitectura súper-modular es capaz de integrar contundentemente las influencias de la información pragmática con los procesamientos y mecanismos implicados en efectuar la tarea relevante, lo que hace justicia al aspecto indeterminado de cada contexto.

Explicar la propuesta completa del autor respecto al entendimiento del lenguaje natural sale de los alcances de la presente tesis, sin embargo, para mis propósitos, es importante notar que es una propuesta poderosa capaz de brindar una explicación de la arquitectura

cognitiva representacional que revela los procesamientos funcionalistas involucrados en la cognición moral, por lo que en adelante explico brevemente los planteamientos que dan lugar a su propuesta.

En primer lugar, García-Ramírez postula la competencia moral, en analogía con la competencia lingüística, como una capacidad súper-modular. En sus palabras: “la competencia moral es un producto de la cognición de orden superior, que involucra múltiples mecanismos cognitivos. Como tal, puede considerarse una habilidad cognitiva súper-modular en el mismo sentido que el lenguaje natural.” (García-Ramírez, 2019, en borrado, p. 268).

La capacidad súper-modular se caracteriza por ser una capacidad *abierta* que interactúa con otras capacidades cognitivas, en donde las interacciones sociales y el entendimiento mentalista de otros es primordial. Para comprender la característica de ser una capacidad *abierta*, primero debo hacer explícito el punto de vista que propone el autor para explicar la competencia lingüística; a este punto de vista el autor le llama *composicionalidad abierta*.

En breve, la *composicionalidad abierta* se basa en la noción de *composicionalidad* que los teóricos tradicionales del lenguaje han empleado para explicar cómo los humanos entienden el lenguaje natural, de acuerdo con la cual el entendimiento del lenguaje se da en virtud de poseer una maquinaria cognitiva especializada (compuesta por léxico, sintaxis y semántica), y ese lenguaje natural funciona por asignar significado a *cualquier* expresión compleja mediante una función composicional algorítmica (García-Ramírez, 2019: IX, las cursivas son mías).

Frente a esta concepción, el autor reconoce que la *composicionalidad abierta* es un enfoque *composicional* en sí mismo, pero con una diferencia mínima que deriva en grandes consecuencias, puesto que la *composicionalidad abierta* acepta que el significado de expresiones complejas pueda ser determinado por algoritmos dirigidos sintácticamente, lo que rechaza es que así sea para *todas* las expresiones complejas posibles y para *todos* sus usos, por lo que el enfoque de *composicionalidad abierta* permite procedimientos alternativos de determinación de significado, tales como el uso de estrategias heurísticas rápidas y frugales (ver *Ibid.*, p. XII).

Por consiguiente, se trata de resaltar el aspecto de ser una capacidad *abierta*, pues es *abierta* por ser capaz de interactuar con dominios, mecanismos y procesamientos

cognitivos que no son específicos al lenguaje o a la moral (respectivamente), así como también, se beneficia de procedimientos alternativos tales como razonamientos heurísticos. En esta línea, el autor encuentra apoyo empírico para sugerir que tanto la competencia lingüística como la competencia moral son competencias *abiertas* a influencias provenientes de diversos procesamientos generales y no específicos a su dominio; influencias que desempeñan roles sustanciales para los propósitos relevantes de la tarea.

Así, el planteamiento es que “el desarrollo moral resulta de la interacción entre varios mecanismos cognitivos de dominio general. Sorprendentemente, un buen número de estos parecen ser centrales tanto para el lenguaje como para la moralidad, incluida la comprensión de los estados mentales, la capacidad de identificar intenciones, participar en la atención conjunta y usar el razonamiento práctico” (2019, en borrador, p.267).

Una condición a considerar es que aquella información o procesamiento relevante que no pertenece específicamente al lenguaje o a la moral, no son elementos extras que puedan adherirse al algoritmo en función, sino que, más bien, son *influencias* que intervienen e interactúan con la capacidad en cuestión, y esto es posible debido a su magnífica característica de ser una capacidad *abierta*, o bien, un sistema representacional *abierto*.

Las heurísticas rápidas y frugales son estrategias de toma de decisiones útiles para la resolución de problemas específicos en contextos particulares, las cuales son *rápidas* por ser cognitivamente económicas y *frugales* por tomar en cuenta sólo partes de la información disponible. Así, las heurísticas se definen por ser estrategias que permiten encontrar la mejor decisión en relación con una situación particular, por lo que su éxito depende de condiciones externas y del entorno (2019, pp. 100, 102). Con esta característica, no sólo se supera plausiblemente el problema de que una arquitectura que acepta módulos sea incapaz de explicar los constantes reajustes que el medio exige, sino que también, hace justicia al aspecto indeterminado de los contextos particulares y expande las posibilidades de conocimiento. De esta manera, ahora se puede ver más claramente cómo pueden ser superados los desafíos de la modularidad tradicional, los cuales abordo brevemente a continuación.

Entonces, la arquitectura súper-modular es capaz de explicar las bases innatas y con ello defender el aspecto universal, pero a su vez, soporta las estrategias heurísticas externamente determinadas y el involucramiento de información pragmática, con lo que el NMI explicaría los recursos cognitivos empleados en contextos indeterminados, en donde

incluso soporta representaciones no acabadas (en el sentido de que serían superficiales, condescendientes e incompletas, por tanto aquí el concepto no es tan rígido como en las concepciones tradicionales)⁴¹, lo que permite dar cuenta de las demandas moralmente relevantes en contextos particulares. Además, explica la inmensa variabilidad individual y cultural.

En torno a ello, el postulado del autor es que el NMI es la base innegociable de la moral, pero la interacción social y el entorno la desarrollan y la complejizan.

2.5.4. Problemas con el modularismo tradicional

Veamos cuáles son los problemas que resuelve la arquitectura súper-modular respecto al modularismo tradicional, según mi interpretación:

- 1) La primera cuestión que surge es ¿de qué tipo de modularidad estamos hablando?, pues debido a que no hay un módulo de la moral como tal, entonces se deben buscar las bases cognitivas que den pie al surgimiento de la misma.

Para dar respuesta a ello, es necesario recurrir al sistema súper-modular altamente interactivo, que por su naturaleza permite la interacción entre diversos mecanismos, procedimientos etc. no específicos a la moral, los cuales, trabajando en coordinación, ejecutan exitosamente la tarea relevante.

- 2) El modularismo concibe el desarrollo en términos puramente madurativos, lo que implica que las diferencias que se contemplan entre el niño y el adulto respecto a la facultad en cuestión, sólo son de carácter *cuantitativo* y no *cualitativo*, es decir, desde este punto de vista el infante posee los mismos dispositivos que el adulto, pero éstos aún no están completamente desarrollados por lo que no son enteramente funcionales.

Esta cuestión ha sido desafiada empíricamente por la psicología del desarrollo, pues los estudios sugieren que los cambios en la cognición moral son tanto *cualitativos* como

⁴¹ Por ejemplo, Ferreira y sus colegas argumentan que el sistema de comprensión del lenguaje crea representaciones sintácticas y semánticas que son simplemente "suficientemente buenas" (SB) dada la tarea que el sujeto comprensor necesita realizar. Las representaciones SB contrastan con aquellas que son detalladas, completas y precisas con respecto al *input* (Ferreira y Patson, 2007). En sentido concomitante, se ha mostrado que el procesamiento del lenguaje es a menudo con representaciones superficiales, condescendientes e incompletas como interpretaciones (García-Ramírez, 2019, p.98).

cuantitativos, y la arquitectura súper-modular es capaz de soportarlos por permitir que la moral esté sujeta a procesos de desarrollo de cambio y no solo de maduración.

- 3) Debido a la especialización de los módulos que se concibe en el modularismo tradicional, las posibilidades de conocimiento (conceptos, creencias, etc.) que podemos desarrollar son limitadas y están previamente condicionadas.

La evidencia empírica muestra que no hay un módulo moral de dominio-específico como tal y en su lugar sugiere que la misma emerge de la interacción entre distintos mecanismos y procesamientos, tales como: Teoría de la Mente (ToM), entendimiento intencional, entendimiento emocional, principios morales pre-lingüísticos y empatía, por lo que requerimos un sistema representacional *abierto* capaz de soportar las interacciones entre estos múltiples mecanismos y procesamientos de información involucrados en la cognición moral, los cuales además son susceptibles de fuertes influencias del entorno, lo que da lugar a que las posibilidades de conocimiento se expandan.

- 4) El aprendizaje está ligado a dotaciones prefijadas, lo que implica que la experiencia tenga un papel mínimo de mero *desencadenante* y el aprendizaje sea poco flexible.

De nuevo, dado que no hay un módulo para la moralidad, se requiere de un sistema capaz de beneficiarse de la interacción entre diversos mecanismos y procesamientos de información, en donde se involucre de manera sustancial la información pragmática y el razonamiento heurístico. Una vez que contamos con la arquitectura súper-modular que es capaz de hacerlo, el aprendizaje se amplía y se ajusta a las necesidades que cada contexto particular exija.

- 5) El modularismo tradicional apela a códigos prefijados aislados del contexto socio-cultural, lo que resulta en una concepción de información que no se corresponde ni con la ambigüedad ni con el desorden asociados a los procesos naturales del conocimiento que tiene que ver con la elaboración y el uso de significados compartidos. En otras palabras, las reglas comunes a todos los sistemas de información no cubren los procesos desordenados, ambiguos y sensibles al contexto de la creación de significados compartidos.

En este punto, es importante resaltar que la propuesta súper-modular es capaz de explicar los procesamientos cognitivos que involucran la información pragmática en toda regla (en donde se prima el uso del razonamiento heurístico rápido y frugal), por lo que el contexto

socio-cultural influye de manera sustancial a los procesamientos en cuestión. Sin embargo, se verá que una consecuencia de mi propuesta, que otorga autonomía al nivel supra-personal, deriva en una explicación de significados *compartidos* que también tienen que ver con una metodología *relacional*, como aquellos que son co-creados en la interacción en marcha. En otras palabras, desde mi propuesta, los significados *compartidos* también emergen y se explican desde su aspecto *relacional*, el cual abordo en el capítulo 3.

Para fines de claridad, cabría preguntar ¿cuál es la diferencia de los significados compartidos según cada enfoque? los fundamentos de cada enfoque señalan que, desde el enactivismo, los significados surgen de la interacción, por lo que los mismos son co-creados por los participantes de la interacción, mientras que, desde el computacionalismo, a partir de las limitaciones prácticas del contexto, el sujeto emplea diversas estrategias para arribar a la mejor interpretación mediante razonamiento heurístico. Así, la respuesta radica en que, bajo la concepción enactiva, los significados compartidos *emergen* en el proceso de la interacción (de manera relacional) con lo que también se destaca la importancia de la experiencia en la generación de significado, y, en la concepción computacionalista, son representados, ejecutados mediante procesos de información y favorecidos por el uso del razonamiento heurístico.Cuál es la intervención de cada uno de ellos en contextos moralmente relevantes, es una cuestión empírica que aún está por determinarse mediante investigaciones adicionales, sin embargo, lo relevante aquí es notar que bajo estas perspectivas se supera plausiblemente la concepción restringida del modularismo tradicional.

- 6) La objeción que hacen Currie y Sterelny (2000, p. 149) a la cognición social modularista es que la atribución mental no puede surgir de mecanismos cognitivos informacionalmente encapsulados debido a que se requiere de la información contextual para efectuarla.

Esta objeción se supera de manera natural por la capacidad altamente interactiva de la súper-modularidad, en donde hay influencias sustanciales entre los diversos mecanismos, procesamientos, información pragmática y estrategias de razonamiento heurístico, lo que permite involucrar la información contextual de manera relevante para el éxito de la tarea en cuestión.

En suma, argumento que, para lograr superar estas cuestiones, debemos recurrir a una multi-modularidad que implique interacciones entre lo modular y lo no modular, entre lo

específico de dominio y lo de dominio general, así como también, que soporte las intervenciones de otras influencias y que sea capaz de involucrar información pragmática y una suerte de razonamiento heurístico externamente determinado. En particular, requerimos de un sistema representacional *abierto*, el cual es otorgado por la *súper-modularidad*.

Ahora veamos el argumento y el sustento empírico del NMI.

2.5.5. Argumento del NMI

Comienzo con una cita de Hamlin:

"De hecho, una comprensión precisa del desarrollo humano requiere tanto una descripción del estado inicial como una imagen precisa de los impulsos que cambian a partir de ahí". (Hamlin, 2015, p. 22).

Esta cita apunta a una necesidad por la descripción del estado inicial del desarrollo humano, lo cual, en el campo del desarrollo moral, motiva a analizar el estado moral de los infantes pre-lingüísticos. Sin embargo, hasta ahora, los estudios psicológicos al respecto habían excluido a los infantes de sus experimentos, concibiendo que el desarrollo moral se compone de cambios *cualitativos* más que de cambios *cuantitativos* (Hamlin, 2015, pp. 5-6). En otras palabras, se concibe al infante como completamente diferente al adulto moral, ya sea como *amoral* -sin sentido moral- o como *inmoral* -que posee un sentido moral opuesto al de los adultos- (Hamlin, 2013, p. 186).

Aquí, es relevante notar que en la psicología del desarrollo ha prevalecido el supuesto de que el desarrollo moral se efectúa mediante cambios *cualitativos*, mientras que en el modularismo tradicional se plantea que es mediante cambios *cuantitativos*. La diferencia es que, desde lo *cualitativo*, el niño sería enteramente distinto al adulto moral, mientras que, desde lo *cuantitativo*, el infante ya cuenta con los mismos recursos que el adulto, pero éstos no están aún desarrollados y por tanto no son completamente funcionales. En la actualidad, la evidencia empírica sugiere que hay cambios tanto *cuantitativos* como *cualitativos* en el desarrollo de la cognición moral.

Puntualmente, estos debates conciernen a la idea de que los infantes nacen o como *tabula rasa* y que todos los aspectos morales se adquieren mediante el aprendizaje, la socialización y la experiencia después del nacimiento, o bien, que nacen con mentes pre-cableadas con contenido moral específico. Esto da lugar a la discusión acerca de si los

infantes humanos nacen como *tabula rasa* moral, o bien, con ciertas predisposiciones morales, mismo que nos remonta al debate entre empiristas y racionalistas.

La propuesta que defiendo es que al menos *algunos* aspectos de la moralidad son favorecidos por predisposiciones naturales basadas-en principios, aspectos que toman la forma de un cuerpo de información sub-personal que nos orienta normativamente hacia nociones universales de lo correcto y lo incorrecto⁴². Dicho conocimiento o cuerpo de información es guiado por principios morales (no aprendidos) que son parte de nuestro aparataje cognitivo, por tanto, son comunes entre los humanos. Pero a su vez, sostengo que el aprendizaje, la socialización y la experiencia tienen un papel irrefutable en el desarrollo moral y con ello considero que la *autonomía* de los aspectos supra-personales deben ganar un lugar en el estudio de la cognición moral, en tanto que formarían parte de los cambios cualitativos más que cuantitativos. Para ser clara, el papel crucial del aprendizaje y de la socialización en el desarrollo del infante es defendido tanto por el computacionalismo como por el enactivismo, pero la diferencia en mi propuesta es otorgarle *autonomía* al nivel supra-personal, con lo que pretendo destacar el aspecto relacional y sus implicaciones. Esto intenta reforzar la idea de que los procesos de desarrollo son influenciados y moldeados por dinámicas *relacionales* más amplias.

Por supuesto no se trata de trivializar la discusión y defender un innatismo radical y/o un empirismo radical, pues ninguna de las dos perspectivas lo haría, de hecho, al parecer, en la actualidad no hay versiones del todo *puristas* al respecto, más bien, intento destacar las contribuciones que cada marco ofrece a partir de un diálogo iluminador entre ellos. Este enfoque integrador es consistente con la tendencia contemporánea en la investigación, que a menudo busca sintetizar diversas teorías y enfoques para obtener una comprensión más abarcadora de los fenómenos complejos, tal como el desarrollo moral.

En esta línea, se debe reconocer y destacar que en ambas perspectivas se habla del agente como un *participante* en el proceso de aprendizaje, en tanto que en el computacionalismo súper-modular el infante es un *aprendiz activo* que constantemente busca y mantiene interacciones sociales con agentes competentes (García-Ramírez, 2019, p.67). Frente a ello, la diferencia radica en que, desde el computacionalismo, el recurso habilitador es la representación, mientras que, desde el enactivismo, se aboga por una comprensión más

⁴² Es conveniente recordar que los límites entre lo innato y lo aprendido son borrosas, de hecho, hay autores como Narváz (2014) que sostienen que esta división es una falsa dicotomía, sin embargo, considero que hay bases morales que pueden considerarse innatas por su temprana aparición en la ontogenia humana.

situada y relacional, incorporando la interacción directa como un componente esencial de la cognición y el aprendizaje, en donde hay un co-habilitamiento agente-ambiente que contribuye al desempeño cognitivo. Así, ambas perspectivas ofrecen distintas formas de entender cómo los agentes *participan* y *aprenden* en su entorno.

2.5.6. ¿De qué se constituye el NMI?

Como he mencionado, se plantea que los elementos que constituyen principalmente al NMI son: Teoría de la Mente (ToM), entendimiento intencional, entendimiento emocional, principios morales pre-lingüísticos y empatía⁴³.

En particular, García-Ramírez⁴⁴ (2019) se basa en el criterio de Carey⁴⁵ (1997) para identificar principios cognitivos universales presentes en infantes pre-lingüísticos y distinguirlos de aquéllos que sean productos de construcciones culturales. Con esto, se han identificado conceptos o principios morales que son comunes a los individuos humanos entre las distintas poblaciones, lo que el autor llama "principios morales pre-lingüísticos" (García-Ramírez, 2019, en borrador, p. 251).

Los principios morales pre-lingüísticos involucrarían principios tales como: *el sufrimiento es malo; ayudar a los demás es bueno; obstaculizar a los demás es malo; cada uno debe recibir lo que se merece; y es bueno ser empático.* (*Ibid.*, p. 252). La determinación de qué otros principios morales pre-lingüísticos estarían involucrados en el NMI, es una cuestión empírica.

En suma, el planteamiento es que la cognición moral está constituida por un aparato cognitivo dotado evolutivamente que cambia conforme al desarrollo, el cual incluye

⁴³ Se verá en el capítulo 3 que también requerimos de la "empatía asimétrica" del enactivismo.

⁴⁴ Cabe aclarar que el argumento del autor García-Ramírez se enfoca en cuestiones metaéticas, defendiendo principalmente un realismo naturalista, en donde logra hacer una síntesis de lo normativo y lo empírico con la propuesta y evidencia del núcleo moral innato. No obstante, yo me enfoco esencialmente en la evidencia empírica y sus planteamientos para argumentar a favor de un tipo de conocimiento moral general (dotado parcialmente por la evolución) en el marco de la filosofía de las ciencias cognitivas, así como también, advierto sus implicaciones normativas, mas no abordo directamente cuestiones metaéticas.

⁴⁵ Elizabeth Spelke y Susan Carey, psicólogas cognitivas y del desarrollo, han argumentado que la mente humana posee varios sistemas cognitivos denominados "sistemas núcleo de conocimiento", los cuales son típicos de especies, fuertemente canalizados en el desarrollo y ya son evidenciados en edades muy tempranas. El caso de los "sistemas núcleo" es principalmente empírico y consiste en mostrar que existen formas particulares de entender el mundo, retenidas en la edad adulta, pero a su vez distintas del entendimiento completo del adulto, las cuales aparecen demasiado temprano en el desarrollo (Horst, 2016, p. 49).

mecanismos cognitivos como habilidades de evaluación social, principios morales prelingüísticos, entendimiento intencional y emocional de los demás, sensibilidad social y sintonía empática temprana. Desde edades muy tempranas el ser humano muestra aspectos de estos procesamientos cognitivos que favorecen el funcionamiento y desarrollo de la cognición moral. A continuación, presento la evidencia empírica que lo apunala.

2.5.7. Evidencia empírica

- Facultades morales

Para efectuar el análisis de las afirmaciones nativistas morales, como primer paso, presento los planteamientos de Hamlin (2013; 2015), los cuales sugieren que los infantes presentan predisposiciones hacia tres características: i) conducta pro-social, ii) comprensión moral y iii) evaluación moral.

Tomando esto como base, me enfoco en el análisis de la acción moral y del juicio moral, respectivamente, lo que efectúo de la siguiente manera:

1) Respecto al comportamiento/acción moral ==> lo rastreo a partir de la evidencia de (i) conducta pro-social, la cual se fundamenta en la empatía. Se debe tomar en cuenta que en esta línea de evidencia el comportamiento moral fue raro en el primer año de edad, pero incrementa en el segundo año (ver Davidov *et al*, 2011).

2) Respecto a los juicios morales ==> lo rastreo a partir de la evidencia de (ii) comprensión moral y (iii) evaluación moral⁴⁶.

Así, pues, este marco explicativo intenta mostrar que los infantes humanos presentan predisposición natural a la empatía, a la pro-socialidad y a la evaluación socio-moral, lo cual, junto con la experiencia, la socialización y la construcción cognitiva posterior funcionan como elementos que co-participan en el desarrollo moral.

Comienzo con la evidencia que apunala las afirmaciones innatistas del comportamiento moral.

- Evidencia para comportamiento moral

⁴⁶ A la capacidad para efectuar juicios morales se le atribuye como condición necesaria (y no contingente) las capacidades de Teoría de la Mente -ToM- (Monsó, 2017), lo que refuerza el argumento visto en el origen cognitivo del juicio moral en los discursos evolutivos.

Las edades de los infantes que abordo en este apartado incluyen desde las 2 semanas de nacidos hasta los 24 meses de edad, y los elementos que subyacen a la conducta moral son principalmente: la empatía, la sensibilidad al sufrimiento de otros y la sensibilidad a la pertenencia de grupo. Además, al parecer, las acciones pro-sociales son intrínsecamente motivadas, al menos, de manera parcial. Asimismo, se verá que hay una fuerte correlación entre la preocupación empática y los actos pro-sociales activos. La finalidad en esta sección es mostrar que la evidencia empírica favorece notablemente el argumento del uso de los principios morales prelingüísticos y por tanto del conocimiento moral general en el comportamiento moral.

Ahora bien, los comportamientos pro-sociales se refieren a aquellas acciones voluntarias destinadas específicamente a beneficiar o a mejorar el bienestar de *otro* individuo o grupo de individuos (Knafo-Noam, 2016), por lo que es razonable relacionar los comportamientos pro-sociales con comportamientos morales⁴⁷.

Como punto de partida, Hamlin cuestiona si los infantes muestran predisposiciones hacia comportamientos morales y lo explora analizando la empatía, en primer lugar, puesto que, es debido a la empatía que las personas tienden a preocuparse por las necesidades de los demás. Esto claramente es una tendencia o impulso dirigido a otros, y la autora define la empatía como "la tendencia a percatarse y a compartir los estados afectivos de los demás" (Hamlin, 2015). Revisemos la evidencia empírica concerniente al comportamiento moral:

Experimentalmente se ha encontrado que los recién nacidos ya muestran respuestas empáticas respondiendo al llanto de otros neonatos. Esto puede explicarse como *contagio* de llanto, lo cual representaría una forma rudimentaria de empatía. En particular, la evidencia indica que los infantes neonatos (2 semanas de edad) ya presentan empatía y cierta sensibilidad a la angustia ajena (Martin y Clark, 1982), pues muestran empatía por otros recién nacidos, siendo capaces de distinguir entre su propio llanto y el de otros (sensibilidad social), entre infantes pequeños y niños mayores, e incluso entre congéneres y chimpancés. Esto parece sugerir que hay reacción empática temprana ante la angustia de otros, lo que es consistente con el argumento evolutivo.

Asimismo, en los primeros 6 meses de edad los infantes muestran sensibilidad a la pertenencia de grupo (incluso, desde los 3 meses) y sensibilidad al sufrimiento de otros

⁴⁷ Más adelante se verá que un criterio para distinguir lo moral de lo social es mediante la distinción intenciones-resultados.

(Hamlin, 2012), lo que indica que los infantes ya presentan cierta sensibilidad social que pertenece al constructo de la moralidad⁴⁸.

A los 8-10 meses de edad, los infantes muestran *preocupación empática* (como respuesta emocional dirigida a otros), la cual, a los 18 meses de edad se correlaciona con conductas pro-sociales activas, por ejemplo, en paralelo a la preocupación empática, los infantes muestran comportamientos reconfortantes hacia otros (cuando son físicamente capaces), tales como dar abrazos, palmadas y objetos reconfortantes (Hamlin, 2015, p.10).

No obstante, entre los 6 y 14 meses de edad parece suceder algo importante, pues si los infantes no muestran preocupación empática en esta edad, muestran mayores niveles de comportamiento antisocial en la niñez media y en la adolescencia (Hamlin, 2015). Cabe mencionar que esto no contrapone al carácter universal, dado que, si no presentan tal preocupación empática, estaríamos hablando de infantes no neurotípicos o con rasgos antisociales.

A los 12 meses de edad muestran conducta pro-social como buenos comunicadores; señalan para mostrar la ubicación de algo que saben que alguien está buscando (Liszkowski, Carpenter y Tomasello, 2008) y también para advertir a un adulto sobre los resultados de acciones que ellos pueden prever, pero el adulto no (Knudsen y Liszkowski, 2013).

Entre los 14 y 18 meses de edad, los infantes ayudan a lograr objetivos (Hamlin, Tomasello). Los niños pequeños comienzan a ayudar con las actividades diarias del hogar (Rheingold, 1982) y ayudan a los adultos a lograr sus metas instrumentales como recoger objetos caídos o apilar libros (Warneken y Tomasello, 2006). Por otra parte, los infantes a esta edad están dispuestos a compartir recursos, pero sólo cuando se les pide que lo hagan, pues la donación espontánea ocurre, aunque rara vez, hasta el segundo año de vida.

A los 18 meses de edad la preocupación empática (presentada a los 8-10 meses) se complejiza y ahora se relaciona con conductas reconfortantes activas (Hamlin, 2013, p. 187), las cuales incluyen consolar a los que están en angustia, pero también, a esta edad ya ayudan a otros de distintas maneras. La idea es que, en cuanto los infantes son

⁴⁸ En paralelo, en el enactivismo también se defiende la sensibilidad a la pertenencia de grupo y al sufrimiento de otros a muy temprana edad (Satne, 2018), la diferencia es que Hamlin (2012) sugiere que la sensibilidad a las intenciones de otros es anterior a la sensibilidad al sufrimiento de otros.

físicamente capaces, comienzan a complementar sus respuestas emocionales o preocupación empática con diversos *comportamientos pro-sociales activos*, tales como consolar a otros que están angustiados, ayudar a otros a lograr metas, informar a otros sobre cosas que deberían saber y a compartir sus propios recursos (Hamlin, 2013, p. 187). (Evidencia para conducta pro-social ver Dunfield, Kuhlmeier, O'Connell y Kelley, 2011; Eisenberg, Fabes y Spinrad, 2006; Warneken y Tomasello, 2009).

Asimismo, a esta edad los infantes se muestran reacios a dañar a las víctimas. Aquí, cabe mencionar que respecto a la conducta pro-social "moral", parece que, cuando incluye una *víctima*, tal conducta puede presumirse como claramente moral. Hay evidencia que muestra que los niños pequeños (16-21 meses) evitan tomar golosinas de aquéllos que han sido obstaculizados anteriormente por otros, lo que sugiere que los infantes se sienten reacios a dañar a otros que han sido dañados, *i.e.* se muestran reacios a dañar a las víctimas (Hamlin, 2015, p. 17).

Por otra parte, Tomasello (2006) plantea que a los 18-24 meses de edad los infantes se involucran en actos de cooperación *compartida*. Warneken y Tomasello (2006) analizaron el comportamiento de niños entre 18 y 24 meses de edad de la siguiente manera: dos agentes (el niño y el experimentador) deben desempeñar roles complementarios para lograr un objetivo conjunto, por ejemplo, tirar de los extremos opuestos de un tubo para recuperar calcomanías que están ocultas en el interior, o bien, colocar una pelota en un tubo y atraparla del otro lado. Se mostró que los niños de 18 meses ya pasan exitosamente este tipo de tareas de resolución de problemas en conjunto. Asimismo, se observó que este tipo de actividad se realiza de manera consistente y espontánea a los 18-24 meses de edad y, con el tiempo, los infantes se vuelven cada vez más hábiles y expertos. Además, al llevar a cabo estos experimentos, los investigadores encontraron que los niños mayores de 18 meses de edad protestan constantemente (verbal o no verbalmente) si el adulto se desconecta antes de que se complete la actividad compartida y, significativamente, que tales protestas aún surgen incluso si la actividad se completa con éxito y la meta es alcanzada. Los autores también observaron una tendencia en los niños a repetir la acción incluso cuando ésta se completó con éxito. En su opinión, esto muestra que, al hacerlo, los niños no persiguen un objetivo individual (por ejemplo, conseguir la pelota), sino que están interesados en la actividad *compartida* como tal. De manera importante, Tomasello explica esta conducta basada en capacidades ToM.

En adición, la evidencia empírica sugiere que el comportamiento pro-social es *intrínsecamente motivado*, pues, parece que la obtención de premios o castigos después de comportarse pro-socialmente no se vincula directamente con motivar (o no) a los niños a efectuarlo en adelante. (Hamlin, 2015).

Al respecto, se ha mostrado que a los 20 meses de edad los infantes que son premiados después de ayudar, ayudan menos en el futuro en comparación con los que no fueron premiados. Esto es relevante debido a que la reducción de un comportamiento después de un refuerzo externo es un sello distintivo de los actos que se consideran motivados intrínsecamente (Hamlin, 2015, p. 12).

Asimismo, a los 20-23 meses de edad los niños se muestran más contentos y excitados en interacciones en las que "dan" que en interacciones en las que "no dan" y, más, cuando dan sus propias golosinas. Esto implica un "dar" o una "donación" con costo para el infante, lo cual no podría ser explicado como un acto meramente social, sino que se sugiere que esto representa un acto pro-social real y no que el infante sólo busque una interacción social bruta (Hamlin, 2015). Sin embargo, se debe resaltar que, en general, un "dar" espontáneo o una donación espontánea por parte de los infantes ocurre, aunque rara vez, hasta el segundo año de vida.

En esta línea, Tomasello ha sugerido que a los 24 meses de edad los infantes están motivados pro-socialmente de manera intrínseca (en oposición a una motivación puramente social o extrínseca), pues ya ayudan a un experimentador desconocido que lo necesite, en presencia o ausencia de los padres.

- **¿Qué criterio ayuda a distinguir entre lo social y lo moral?**

Hay una manera clásica de distinguirlos, pues desde Piaget (1932) se sugirió un criterio para tal diferenciación, a saber, "la distinción entre intenciones *versus* resultados". Con base en ello, la evidencia sugiere que a los 10 meses de edad los infantes prefieren a los que ayudan *intencionalmente* sobre aquéllos que ayudan *accidentalmente*, y a los que obstaculizan *accidentalmente* sobre los que obstaculizan *intencionalmente*.

Sin embargo, sorprendentemente se ha visto que a los 8 meses de edad los infantes ya priman las intenciones sobre los resultados, lo que está involucrado en sus evaluaciones y juicios morales (Hamlin, 2012). Se debe destacar que a esta edad (8 meses), los estados mentales ya son cruciales para las elecciones sociales de los niños (Hamlin, 2015, p.19).

Esto es relevante debido a que los estados mentales son críticos para los juicios morales tanto de los niños como de los adultos.

Según Hamlin (2012), de hecho, primero aparece la sensibilidad a las *intenciones* de otros, más que a sus estados emocionales, con lo que sugiere que no hay respuesta al sufrimiento de otros antes de que haya respuesta a las intenciones de otros. Entonces, en este marco empírico, la atención a las intenciones de otros precede a la sensibilidad al sufrimiento de otros (que se presenta en los primeros 6 meses de edad).

Hasta aquí la evidencia empírica respecto a la conducta moral, ahora revisemos la evidencia pertinente a los juicios morales.

- **Evidencia para juicios morales en infantes pre-lingüísticos**

Para comenzar el análisis de los aspectos relevantes concernientes a los juicios morales en pre-lingüísticos, cabe resaltar dos características principales de los mismos:

- 1) Los juicios morales son evaluaciones (sociales).
- 2) Tales evaluaciones son hechas esencialmente por un *observador* no implicado directamente en la situación (Hamlin, 2015). Por ello, se le atribuye como condición necesaria (y no contingente) las capacidades de ToM, *i.e.* se asocia con una perspectiva socio-cognitiva de tercera persona.

Ahora bien, la pregunta crucial es saber si a los infantes les importa *qué hace qué a quién*, y la evidencia arroja resultados que sugieren que sí muestran interés de ello desde los 6-10 meses de edad y, de hecho, como se verá más adelante, los estudios se extendieron, incluso, a infantes de 3 meses de edad.

A continuación, presento tal evidencia, la cual apuntala fuertemente las capacidades evaluativas de los infantes pre-lingüísticos.

Para analizar si a los bebés les importa *qué hace qué a quién*, se realizan experimentos llamados “eventos de ayudante-obstaculizador”, que consisten en lo siguiente:

En el estudio original, Hamlin, Wynn y Bloom (2007) mostraron a infantes de 6 a 10 meses de edad una secuencia de eventos en los que una figurilla, un “escalador”, ascendía por una pendiente en tres ocasiones sucesivas. En las dos primeras ocasiones llegaba a la mitad de la pendiente y luego volvía a bajar hasta la parte inferior de ella. En la tercera

ocasión, otra figurilla entró en escena cuando el “escalador” estaba a la mitad del camino. En algunas pruebas, la segunda figurilla era un “ayudante” que hacía contacto con el “escalador” y juntos se movían hasta la parte superior de la pendiente. En otras pruebas, la segunda figurilla era un “obstaculizador” que hacía contacto con el “escalador” y juntos se movían hasta la parte inferior de la pendiente.

La metodología que se emplea en el experimento es un “paradigma de elección” para determinar la *preferencia* (los bebés alcanzan lo que prefieren) y un “paradigma de violación de expectativas” para determinar la *sorpresa* (los bebés miran más tiempo los eventos sorprendentes). (García-Ramírez, 2019, en borrador, p.252)

Los resultados mostraron que en pruebas de *preferencia* los infantes eran más propensos a alcanzar al “ayudante” más que al “obstaculizador” cuando los dos se presentaban uno al lado del otro y ellos se sorprendían más (como lo indica el tiempo de mirada) cuando el “escalador” se movía hacia el “obstaculizador” que cuando se movía hacia el “ayudante”.

Esto sugiere que los infantes de 6 a 10 meses de edad ya son capaces de efectuar evaluaciones sociales que involucran valores positivos y negativos, y con ello, el cálculo de intenciones, evaluación de conductas y la distinción entre agentes que *ayudan* o *cooperan* y agentes que *obstaculizan* (Hamlin, Wynn y Bloom, 2007). Esta capacidad evaluativa compleja parece ser la base de un sistema de cognición moral en desarrollo, dado que en términos adaptativos la cooperación es central al éxito de la especie y para alcanzarla es necesario distinguir entre cooperadores y obstaculizadores.

Ahora bien, estos experimentos se extendieron a infantes de 3 meses de edad y los resultados mostraron sorprendentemente que ellos también prefieren a los que ayudan y no los que obstaculizan, aunque no distinguen entre ayudantes y neutros (Hamlin, 2015, p. 16). Por tanto, se sugiere que se desarrolla primero una aversión a los actos anti-sociales, la cual ya está presente en infantes de 3 meses de edad.

Es relevante mencionar que, previo a presentar tal aversión a los actos anti-sociales, los bebés muestran sensibilidad a la pertenencia de grupo, y no sólo la muestran, sino que la anteponen frente a tal aversión. Hay estudios que muestran que los infantes consistentemente prefieren a los que son similares a ellos (por ejemplo, en idioma o acento) más que a los disímiles, o bien, que prefieren a los que son de su grupo más que a los que no son de su grupo. En paralelo con esta línea de evidencia se realizaron experimentos de

“eventos de ayudante-obstaculizador”; los resultados arrojados mostraron que los infantes en realidad prefieren al que obstaculiza a los que son diferentes a ellos, y no al que ayuda a los que son diferentes a ellos (Hamlin, 2012, p. 170). Esto parece sugerir que previo a la aversión de actos anti-sociales, los bebés desarrollan la sensibilidad a la pertenencia de grupo.

En suma, a los 10 meses de edad prefieren al que ayuda sobre el que obstaculiza y, como se vio, a esta edad ya prefieren al que ayuda intencionalmente sobre el que ayuda accidentalmente, así como también, prefieren al que obstaculiza accidentalmente sobre el que obstaculiza intencionalmente. Estas capacidades evaluativas parecen involucrar ToM y representaciones de 2o orden. Aquí, es relevante recordar que a los 8 meses de edad los infantes priman las intenciones sobre los resultados y los estados metales ya son cruciales para sus elecciones sociales. Incluso, como se vio, los bebés prestan atención a las intenciones de otros antes de presentar sensibilidad al sufrimiento de otros.

Tales capacidades se van complejizando y a los 10-19 meses de edad prefieren al que da al ayudante y al que quita al obstaculizador, lo que parece ser un mecanismo en desarrollo del "enemigo de mi enemigo", pues en ciertos contextos los infantes evalúan positivamente actos anti-sociales tales como el castigo (Hamlin, 2015). A esto vuelvo más adelante.

Ahora bien, respecto a la evaluación de la conducta cooperativa, Hamlin (2013, p. 187) señala que los infantes, al final de su primer año de edad, entienden que los agentes trabajan juntos para alcanzar objetivos compartidos (Henderson y Woodward, 2011); clasifican la *ayuda* para alcanzar objetivos como positiva y su *obstaculización* como negativa (Premack, 1997); y entienden que ser *ayudado u obstaculizado* influye en las preferencias sociales (Kuhlmeier, Wynn y Bloom, 2003). Esto parece mostrar una base sólida para el desarrollo de la comprensión moral.

En consonancia, a los 12-16 meses muestran una comprensión de estados emocionales más compleja y a los 14 meses de edad manifiestan un complejo entendimiento del trabajo colaborativo.

Además, los infantes también muestran ser sensibles a la noción de *justicia*, pues, a los 2.5 años de edad (30 meses), ya esperan que los individuos traten a los otros de manera justa, por ejemplo, dividiendo igualmente los recursos (Geraci y Surian, 2011), siempre que todos

hayan contribuido por igual al esfuerzo cooperativo o estén en el mismo grupo (Sloane, Baillargeon y Premack, 2012, en Hamlin, 2013, p.187).

En sentido concomitante, parece haber un vínculo estrecho entre las expectativas individuales de justicia y los comportamientos pro-sociales de los infantes, pues Schmidt y Sommerville (2011) mostraron que los niños de 15 meses que se involucraron en actos pro-sociales de "dar" (dar el mejor juguete a otro) también miraron por más tiempo las distribuciones desiguales (Hamlin, 2015, p. 21). Y, a los 20 meses de edad ellos miran por más tiempo los actos distributivos desiguales que los iguales, a menos que los receptores de la acción hayan realizado anteriormente un trabajo desigual (Sloane, Baillargeon y Premack, 2012 en: Hamlin, 2015, p..21).

De hecho, infantes de 16 meses de edad ya muestran preferencia por las distribuciones justas que por las injustas, lo que sugiere, según los autores, que el componente evaluativo clave de *justicia* ya está presente en las evaluaciones que efectúan los pre-lingüísticos respecto a los actos distributivos de los demás.

Tras revisar esta evidencia empírica se puede sugerir que las capacidades evaluativas socio-morales de los niños pequeños parecen ser robustas, complejas y altamente mentalistas⁴⁹, por lo que a continuación presento la evidencia que entrelaza tales capacidades con teoría de la mente.

- **Juicio moral y teoría de la mente (ToM)**

Primero, Hamlin, Ullman, Tenenbaum, Goodman y Baker (2013) exploran si los infantes de 10 meses de edad involucran habilidades mentalistas durante el proceso de evaluación social. Los resultados muestran que, en efecto, los infantes pre-lingüísticos toman en cuenta la información mentalista y realizan inferencias en el curso de sus evaluaciones sociales, estimando así el papel que desempeñan los estados mentales en las acciones de otros.

Por otra parte, Kovács y Teglas (2010) presentan un experimento que sugiere que infantes de 7 meses de edad (igual que los adultos) codifican automáticamente las creencias de

⁴⁹ El término "mentalista", "Teoría de la mente" y "lectura de mentes" se usan en sinonimia y refieren a la habilidad de entender, predecir y explicar la conducta de otros mediante la atribución de estados mentales, tal como se vio en la sección de la cognición social computacionalista.

otros y que las creencias de otros tienen efectos similares a las creencias propias, así como también, parece que se tienen en cuenta como representaciones alternativas al entorno.

Ahora bien, hay evidencia robusta que muestra la relación directa entre Teoría de la mente (ToM) y cognición moral. Por ejemplo, Wellman y Miller (2008) argumentan que las perspectivas deónticas y mentalistas se integran mutuamente; tanto el razonamiento de creencia-deseo como el de obligación-permisos son de desarrollo temprano, universales e inseparables en la comprensión de los estados mentales de los demás. Los conceptos básicos para el razonamiento deóntico son los de *obligación* y *permiso*, y éstos incorporan una noción de fondo de la sensibilidad social a la acción intencional, por lo que las consideraciones deónticas juegan un papel crucial en la comprensión y en la explicación del comportamiento intencional.

Infantes de 26 meses de edad ya consideran como "malo" violar las obligaciones morales (por ejemplo, morder a otro niño) y a los 3-4 años de edad ya ven estos actos como obligaciones desobedecidas y, por tanto, merecen un castigo y una evaluación negativa. Además, hay buenas razones para pensar que Teoría de la mente y razonamiento deóntico se desarrollan en paralelo; cuando ToM cambia, alrededor de los 4 años de edad, también lo hace el razonamiento sobre las acciones como conforme (o no) a las reglas y obligaciones (ver Wellman y Miller, 2008).

Por otra parte, algunos estudios sugieren que es necesaria una comprensión coordinada de los estados mentales (ToM) y de los estados emocionales para reconocer cómo es que los demás aprueban o desaprueban nuestras acciones y por tanto para el desarrollo moral (implica razonamiento moral y toma de decisiones). Asimismo, la comprensión de los estados emocionales y mentales de los demás ayuda a los infantes a ver más allá de las necesidades propias y a centrarse en necesidades físicas y psicológicas de los demás. Esto sucede aproximadamente a los 3.5 años de edad (Lane y Wellman, 2010).

Se debe agregar que una cuestión crucial concerniente a los juicios morales es preguntar si los infantes *¿evalúan positivamente actos antisociales* tal como los adultos lo han hecho a lo largo de la historia con el castigo a los infractores?

La evidencia muestra que entre los 10-19 meses de edad los infantes prefieren al que *da* (golosinas) al ayudador y al que *quita* (golosinas) al obstaculizador. Esto sugiere que hay interpretaciones flexibles y sensibles al contexto por parte de los infantes en edades muy

tempranas. En particular, como mencioné anteriormente, esta evidencia muestra que la interpretación de los infantes parece tomar la forma de un mecanismo en desarrollo de "el enemigo de mi enemigo". Según Hamlin, un mecanismo así es muy benéfico, puesto que se puede aplicar a cualquier persona que pueda causar daño, incluidos los obstaculizadores, pero también a los miembros del grupo, amigos de enemigos, etc. (Hamlin, 2013, p. 191).

En sentido concomitante, recordemos que en la sección de "el origen de los juicios morales" se vio que los estudios evolutivos sugieren que su evolución cognitiva surgió por la necesidad de detectar a los *tramposos* para evitar así que se beneficiaran de la vulnerabilidad de los sistemas cooperativos, es decir, la presencia de los tramposos exigió a los individuos desarrollar la capacidad de *leer* otras mentes.

Así, pues, estimo que tras revisar estas investigaciones experimentales, parece que tenemos buenas razones para pensar que los infantes prelingüísticos muestran fuertes predisposiciones morales, al menos para la conducta pro-social, la comprensión moral y la evaluación moral, tal como lo sugiere Hamlin.

Asimismo, la evidencia empírica parece sugerir que los dominios que constituyen la arquitectura cognitiva de la moralidad (o del NMI) son preeminentemente: principios morales prelingüísticos, teoría de la mente, entendimiento intencional, entendimiento emocional y empatía, los cuales interactúan entre sí y funcionan como procesos de información cargados de contenido.

Ahora bien, como es de esperarse, la función de estas capacidades pre-lingüísticas depende en gran medida de su desarrollo y paralelamente sucede un proceso por el cual el contexto y la cultura modifican, enriquecen o suprimen la emergencia de la cognición moral para adquirir un código moral específico. Por ello, me interesa mencionar a continuación algunas consideraciones respecto al desarrollo posterior de la moral.

2.5.8. Algunas anotaciones respecto al desarrollo posterior de la moral

La investigación psicológica ya no explica el desarrollo moral como una secuencia progresiva, estricta e invariable de etapas (como en Kohlberg), sino que, ahora lo explica como un proceso dinámico y cambiante, por lo que existen revisiones y regresiones en el mismo, así como también, se beneficia de la interacción con otros procesamientos cognitivos no morales (tales como ToM, entendimiento de intenciones, sensibilidad al

contexto, empatía, etc.), característica que, como se vio, es posible por la capacidad súper-modular.

Ahora bien, el desarrollo moral se realiza dentro de un contexto cultural en el que el sujeto practica los comportamientos en los que está inmerso y participa de las creencias, valores, sanciones, reglas, motivos y satisfacciones asociados con ellos (Lapsley, 2014), lo cual da lugar a la variabilidad entre los códigos morales relativos a las culturas.

De esta manera, la cognición moral da lugar a enormes variaciones individuales y socioculturales. Respecto a lo individual, por ejemplo, la evidencia disponible sugiere que la relación padre-hijo en edades tempranas (1-2 años de edad) es crucial para el desarrollo de una "consciencia moral" pertinente a normas y valores, es decir, para el desarrollo de un sistema de guía interno que regule el comportamiento del individuo. Kochanska (2002) sugiere que en las relaciones padre-hijo con *orientación mutuamente receptiva* (la cual es positiva), el niño desarrolla tal sistema de guía interno de manera natural, mientras que los niños que carecen de ese tipo de relación con sus padres, lo desarrollan con demasiado esfuerzo, por tanto, les requiere mayor trabajo regular su conducta.

Otro aspecto relevante es la identidad moral, que se vincula principalmente con la calidad de la relación padre-hijo, en la que se ha observado que los infantes de 25-52 meses de edad muestran seguir las reglas maternas sin que se les observe (Kochanska, Koenig, Barry, Kim y Yoon, 2010).

En concreto, existe un amplio cuerpo de trabajo empírico que detalla los resultados de estudios longitudinales acerca del desarrollo moral y su variabilidad individual y cultural, en los que parece haber un desarrollo estable en la niñez, una disminución importante del mismo en la adolescencia y un repunte en la adultez; en los que se muestra de manera contundente la intervención e interacción de los recursos cognitivos mencionados, aunados a rasgos y temperamentos individuales, como también, a factores contextuales y socio-culturales (ej., Kochanska, 2002; Super y Harkness, 1997; Whiting y Edwards, 1988; Knight y Carlo, 2012; de Guzman, Edwards y Carlo, 2005).

Estas investigaciones han representado esfuerzos por avanzar en las explicaciones concernientes a los cambios de desarrollo en la moralidad. En particular, los estudios abordan los rasgos de personalidad y temperamentos individuales, modelos de socialización, enfoques del desarrollo cognitivo y mecanismos vinculados a la cultura, los

cuales se amplían a partir de la comprensión básica que tenemos respecto a la cognición moral, a las emociones y a los comportamientos morales. No obstante, estas líneas de investigación “vienen con complejidades y desafíos adicionales que requieren un discurso más co-constructivo en los diversos subcampos de estudio moralmente relevantes” (Lapsley y Carlo, 2014, p.5), lo que nos autoriza a explorar otros marcos teóricos-metodológicos capaces de complementar las explicaciones e investigaciones disponibles.

Así, pues, mi intención aquí ha sido presentar el amplio cuerpo de trabajo empírico que apuntala las afirmaciones nativistas concernientes a la moral, puesto que, lo referente a la variabilidad cultural, también lo abordo desde una metodología relacional y, en consecuencia, desde un discurso que otorga autonomía al nivel supra-personal, es decir, desde el enactivismo.

En suma, me interesa resaltar las ventajas que presenta otorgar autonomía al nivel supra-personal, así como también, la evidencia que lo sustenta, la cual señala que “los recientes desarrollos conceptuales y empíricos indican consistentemente la necesidad de investigaciones que permitan el estudio de encuentros sociales en tiempo real de una manera verdaderamente interactiva. Esta sugerencia se basa en la premisa de que la cognición social es fundamentalmente diferente cuando interactuamos con otros en lugar de simplemente observarlos” (Schilbach, *et.al.*, 2013).

No obstante, debe quedar claro que respecto a la parte computacionalista de mi marco integrativo, he recurrido a la arquitectura súper-modular debido a que es capaz de soportar lo innato, pero a su vez, de involucrar sustancialmente la información pragmática en sus procesamientos; característica que además de presentar genuinas ventajas ante las explicaciones computacionalistas disponibles, se ajusta bien con la arquitectura cognitiva PI que propongo. Asimismo, se debe subrayar que el NMI explica contundentemente el nivel supra-personal, no obstante, mi intención es resaltar el papel que éste desempeña desde un análisis *relacional*.

Por último, para cerrar esta sección, considero útil exponer la línea cronológica de las capacidades (constituyentes del NMI) que se vieron, esto con el objetivo de identificar claramente lo que ocurre en cada etapa del desarrollo pre-verbal. Por consiguiente, a continuación presento la incorporación de los procesamientos que subyacen al comportamiento moral, al juicio moral y a las capacidades de teoría de la mente.

2.5.9. Línea cronológica del NMI: el desarrollo moral pre-lingüístico

Capacidades constituyentes del NMI:

2 semanas: contagio de llanto como un tipo de empatía rudimentaria.

Previo a los 3 meses: sensibilidad a la pertenencia de grupo. Esto es previo a diferenciar claramente entre ayudante/obstaculizador. No obstante, un dato interesante es que los infantes, respecto a sus disímiles, prefieren a los que les obstaculizan más que a los que les ayudan.

3 meses: prestan atención a las intenciones de otros y presentan aversión a actos antisociales, pues prefieren al ayudante sobre el obstaculizador, pero no distinguen entre ayudante y neutro.

6 meses: ya siguen la mirada, lo que es fundamental para predecir intenciones y presentan sensibilidad al sufrimiento de otros.

7 meses: funcionamiento inicial de ToM, el cual parece ser automático y sin esfuerzo. Además, ya forman representaciones de las creencias de otros.

8 meses: priman intenciones sobre resultados y los estados metales ya son cruciales para sus elecciones sociales.

8-10 meses: preocupación empática cognitiva y afectiva.

9 meses- Tomasello plantea que hay una revolución socio-cognitiva en los infantes, pues muestran toma de perspectiva, atención conjunta, se involucran en situaciones triádicas y de intencionalidad compartida. También muestran comunicación simbólica tal como señalar y hacer gestos.

10 meses: ya son capaces de evaluaciones sociales mentalistas que incluyen representaciones de segundo orden. Esto se ve claramente cuando a esta edad prefieren al que ayuda sobre el que obstaculiza; asimismo, prefieren al que ayuda intencionalmente sobre el que ayuda accidentalmente, y, al que obstaculiza accidentalmente sobre el que obstaculiza intencionalmente. Tales evaluaciones socio-morales involucran ToM.

10-19 meses: prefieren al que *da* al ayudante y al que *quita* al obstaculizador, esto parece favorecer un mecanismo del "enemigo de mi enemigo". También, evalúan positivamente actos anti-sociales como el castigo.

12 meses: conducta pro-social como buenos comunicadores.

Al final del 1er año: evaluación de conducta cooperativa más sofisticada.

12-16 meses: comprensión de estados emocionales más compleja.

14 meses: complejo entendimiento del trabajo colaborativo.

15 meses: hay noción de justicia; parece haber un vínculo entre las expectativas individuales de justicia y los comportamientos pro-sociales de los infantes, pues los que se involucraron en actos de "dar" (dar el mejor juguete a otro) miraron por más tiempo las distribuciones desiguales (ver por más tiempo representa mayor sorpresa, i.e. es un evento no esperado).

16 meses: claramente prefieren las distribuciones justas.

14-18 meses: ayudan a lograr objetivos.

18-24 meses: se involucran en conducta de cooperación compartida.

18 meses: la preocupación empática se complementa con conductas reconfortantes activas: consuelan dando palmadas, abrazos y objetos reconfortantes.

16-21 meses: se muestran reacios a dañar a las víctimas.

20-meses: muestran que las conductas pro-sociales son intrínsecamente motivadas; los que son premiados por conducta pro-social ayudan menos en el futuro.

20-23 meses: se involucran en actos de dar y, además, dan con costo personal.

24 meses: ayudan a otros en presencia o ausencia de los padres.

26 meses: consideran como malo violar las obligaciones morales (ej. morder a otro niño).

30 meses (2.5. años)- entienden retribución, esto enraizado en la habilidad de evaluar acciones morales; esperan que los demás traten a los otros de manera justa y dividen igualmente los recursos si todos cooperaron igual o son del mismo grupo.

3-4 años: ven la violación moral como obligaciones desobedecidas y merecen evaluación negativa y castigo.

4 años: cambia ToM y cambia el razonamiento deóntico (con nociones de obligación-permiso): hay razonamiento sobre las acciones como conforme o no a las reglas y obligaciones.

4-5 años: entienden las emociones del otro con relación a sus metas.

5.5 años- muestran altos niveles de ToM y de entendimiento emocional: ToM para comprender necesidades psicológicas y entendimiento emocional para entender necesidades físicas y materiales.

7 años (y adultos)- entienden las emociones de otros con relación a reglas y consecuencias que pueden tener.

Ahora corresponde ver las implicaciones normativas de la evidencia presentada.

2.6. Normatividad

2.6.1. Propuesta normativa: conocimiento moral general objetivo y sus implicaciones normativas

Como punto de partida, es crucial tener en cuenta que las implicaciones normativas que propongo en mi marco integrativo son a partir de su doble metodología-cognitiva, de sus dos entidades núcleo y de sus distintos niveles de análisis.

En torno a ello, mi propuesta normativa comprende dos niveles: 1) como una estructura sistemática de un conocimiento moral general que es parte de nuestro aparato cognitivo, lo que asegura la objetividad moral, y; 2) como una estructura social que analizo desde una metodología relacional, la cual se instituye en prácticas sociales y se mantiene mediante las interacciones. Si mis consideraciones son correctas, esto permite ofrecer una herramienta evaluativa capaz de generar juicios *situados*, pero anclados a criterios *objetivos*.

En esta sección, me centro en exponer concisamente las implicaciones normativas computacionalistas, pues, para presentar la propuesta completa es necesario revisar las contribuciones enactivas, las cuales presento en el capítulo 3.

Entonces, ¿qué implicaciones normativas derivan del modo-cognitivo computacionalista que se vio, de su metodología individual y de su entidad teórica núcleo que es la *representación*?

Defiendo que la implicación normativa fundamental que deriva de ello es la existencia de un tipo de conocimiento moral general que nos es dotado evolutivamente y como tal es parte de nuestra maquinaria cognitiva. Con ello, se asegura la objetividad moral y se evita la apropiación cultural de los fundamentos morales, lo que es posible por basarlos en propiedades psicológicas que son naturalmente parte de nuestro aparato cognitivo.

Ahora bien, mi postura respecto a las afirmaciones nativistas, es *moderada*, debido a que, si bien, acepto el papel innato y universal de un tipo de conocimiento moral general favorecido por la evolución, también defiendo que ese conocimiento es de alcance limitado, en el sentido de que, por su generalidad, no puede dar cuenta de lo particular⁵⁰, es decir, de extraer los contenidos particulares de las situaciones específicas. Para ello, defiendo que se requiere también de una metodología relacional capaz de develar el carácter *situado* de la moral y como estructura compartida.

En este punto debe quedar claro que el NMI explica los recursos cognitivos empleados en contextos indeterminados (que son guiados principalmente por razonamiento heurístico), en donde incluso soporta representaciones no acabadas (como las representaciones *SB* que se mencionaron en la nota al pie 41), que permiten explicar las demandas moralmente

⁵⁰ Esto es de esperarse, pues los argumentos evolutivos sostienen que lo particular o lo específico a las situaciones no puede ser heredado biológicamente (Darwin citado en Etzelmüller, 2016).

relevantes en contextos particulares. No obstante, como he mencionado, mis objetivos respecto a lo contextual (o al contexto interactivo) son distintos, pues en mi propuesta, requerimos tanto de una explicación de los procesamientos representacionales empleados en la moral, como también del aspecto relacional emergente de las interacciones y de las practicas sociales compartidas.

En particular, planteo que la evolución nos dota naturalmente con principios morales generales que son la *condición de posibilidad* de otros valores morales relativos a las culturas, los cuales, sostengo, pueden también emerger relacionalmente, es decir, en ocasiones, las derivaciones culturales y el contenido-situado de esos principios es determinado relacionalmente a partir del contexto socio-cultural de prácticas compartidas al que pertenece el agente en cuestión; implicación normativa que deriva de mi propuesta integrativa. Así, en mi interpretación, los principios morales innatos son generales, pero sus derivaciones se caracterizan por poseer cierto pluralismo de realizabilidad, es decir, tales principios son de carácter amplio, pues sus derivaciones culturales y su contenido específico lo dicta el contexto y la cultura, al menos, en situaciones problemáticas que conducen a dilemas morales.

Me explico; por ejemplo, el principio general de “ayudar es bueno” nos orienta normativamente y nos motiva a hacerlo, pero en ocasiones, la derivación de su contenido específico obedece en gran parte a señales contextuales y a las necesidades concretas que la situación exige, por lo que presenta matices y diversas formas de realización que se explican a partir de los significados compartidos que emergen de las interacciones concretas, lo que propongo, también debe estudiarse desde una metodología relacional. Así, en diversos casos, la situación, el contexto, la cultura y la comunidad a la que pertenecemos dictaría cómo se realiza específicamente la derivación de dicho principio (considérense ejemplos controversiales como: “la ablación femenina *ayuda* a realzar la lealtad de la mujer con lo que aumenta sus posibilidades de formar una familia”).

Para ser clara, sin duda existen muchos casos que no son controversiales y son aceptados universalmente, tal como lo es “ayudar a una persona de edad avanzada a cruzar la calle es bueno”, no obstante, existen casos evaluativos más controversiales que justo son mi tema de interés, tales como prácticas ritualistas que implican mutilaciones o prácticas médicas controversiales como la eutanasia, etc.

En esta línea, por ejemplo, un dilema ético comúnmente abordado por las teorías normativas es la práctica de la ablación femenina, la cual se analiza desde perspectivas morales que defienden, o, una visión basada en criterios objetivos y por tanto aplicados a *todas* las situaciones, contextos y culturas que generan los mismos juicios morales, o bien, desde perspectivas que defienden una visión relativista y enteramente cultural, lo que implica que carezcan de una base objetiva y se corra el riesgo de la apropiación cultural de la moral.

Al respecto, en mi propuesta considero ejemplos precisos de dicha práctica, los cuales presentan amplia diversidad en los principios, motivos y valores que les originan (trabajo realizado con mi colega Patricia Melchor en: Pescador y Melchor, 2024, en revisión), de lo que se plantea que, para analizarlos moralmente, se requiere evaluar su aspecto situado desde una metodología relacional y con ello dar voz a las agentes involucradas y a las circunstancias específicas que les envuelven. Pero a su vez, precisamos también de una base objetiva capaz de evaluar las prácticas, acciones y culturas opresivas de las que no lo son, así, si mis consideraciones son correctas, la integración de mis dos metodologías (que deriva en la integración de lo objetivo y lo situacional), permite sugerir una herramienta evaluativa capaz de generar juicios *situados* pero anclados a criterios *objetivos*.

Este tema es crucial y lo analizo a profundidad en el siguiente capítulo, debido a que el aspecto *situado* es una consecuencia normativa del enactivismo y de su metodología relacional, lo que debe ser explicado claramente para habilitar mi propuesta.

En suma, la propuesta normativa de doble metodología deriva en: 1) defender la objetividad moral por ser parte de nuestro aparato cognitivo, y, 2) abordar su aspecto situado, cultural y contextual desde una metodología relacional, lo que permite hablar de la moralidad como una estructura social e instituida en la práctica social. Así, propongo que la normatividad moral se puede explicar desde dos aspectos constituyentes; como una estructura sistemática dotada evolutivamente capaz de asegurar la objetividad moral y, también, como una práctica social mantenida relacionamente, lo que devela su aspecto *situado*.

El valor de dicho conocimiento moral general y objetivo es que permite identificar la relación de similitud moral que hay detrás de la vasta diversidad cultural, con lo que es posible distinguir las acciones, prácticas y culturas moralmente opresivas de las que no lo son.

Veamos el argumento condensado.

2.6.2. Argumento normativo condensado

- 1) La evolución nos dota con un tipo de conocimiento moral general; hipótesis que encuentra apoyo en la evidencia empírica de infantes pre-lingüísticos.
- 2) Tal conocimiento moral general es un cuerpo de información que nos guía normativamente en contextos moralmente relevantes, siendo éste un conocimiento psicológico sub-personal que comprende principios morales universales.
- 3) Por tanto, este cuerpo de información funge como herramienta evaluativa objetiva capaz de orientarnos normativamente.

A partir de ahí, definiendo que contamos con una preparación natural para la adquisición de la moralidad, la cual se compone de principios morales prelingüísticos y de otros mecanismos cognitivos que en su interacción le favorecen, pero sus adaptaciones son crucialmente perceptivas y dependen del contexto normativo en el que el agente nace y se desenvuelve.

Basar la moral en principios universales con valor objetivo y transcultural no impide hablar de códigos morales relativos a las culturas, al contrario, da la base objetiva para distinguir lo que es moralmente opresivo de lo que no lo es y así evitar la apropiación cultural de los fundamentos morales.

Pero ¿qué asegura tal objetividad? son los principios morales prelingüísticos los que aseguran una moral objetiva, pues son parte natural de nuestra dotación psicológica. La justificación de la existencia de esos principios morales es que son parte de nuestra naturaleza cognitiva y, como se vio, existe un sólido apoyo empírico que lo sugiere.

Ahora bien, el papel normativo que desempeñan los principios morales pre-lingüísticos es que generan los juicios morales respecto a las nociones universales concernientes a lo correcto-incorrecto, bueno-malo, etc., es decir, regulan y determinan (en sentido amplio) cómo deben llevarse a cabo las conductas y juicios morales, así que regulan nuestros comportamientos y evaluaciones, y determinan lo que es incontrovertidamente aceptable-inaceptable, loable-despreciable, etc., dando lugar así a nuestros juicios morales en general, lo que su vez, constituye la *condición de posibilidad* de otros valores morales relativos a la cultura (mismos que, en parte, se explican *relacionalmente*).

Así, la propuesta normativa que definiendo es clara: el computacionalismo otorga los *fundamentos* morales basados en principios evolutivamente dotados y por tanto objetivos y

universales, y, el enactivismo, por su parte, explica y evalúa relacionamente el nivel supra-personal y con ello el aspecto *situado* de la moral. De ahí, propongo que deriva una herramienta evaluativa capaz de generar juicios *situados* pero anclados a criterios *objetivos*.

Esto es posible debido a que mi propuesta se basa en conceder que cada marco conceptual se enfoca en explicar distintos aspectos de la cognición socio-moral, por lo que su complementación permite explicaciones más abarcadoras, tanto cognitivas como normativas respecto al tema de interés, o al menos así lo defenderé.

Ahora bien, para cerrar esta sección, a continuación presento algunas consecuencias que me interesa tener en mente para dar paso al siguiente capítulo dedicado al enactivismo.

2.6.3. Consecuencias

1. Considero que la propuesta del NMI es parsimoniosa y presenta genuinas ventajas frente a otros enfoques computacionalistas, debido a que es capaz de superar los debates tradicionales que han prevalecido: explica los procesamientos cognitivos del individuo y a su vez reconoce las fuertes influencias que tiene el entorno en ellos; lo que explica en términos de ser computados representacionalmente por el sujeto.
2. En contraste, en mi propuesta integrativa se aborda la moral también como una estructura social (que se instituye en la práctica social y se mantiene en las interacciones), otorgando autonomía a los procesos supra-personales involucrados en la cognición moral, movimiento conceptual que encuentra justificación en un amplio cuerpo de trabajo sociológico.
3. Concretamente, estudiar la moral como una estructura social es para resaltar su dimensión social desde un enfoque relacional, lo que presenta al menos tres ventajas: 1) devela el aspecto *situado* de la moral mediante los significados compartidos que emergen de la interacción, 2) explica la co-constitución agente-entorno o agente-agente, según sea el caso y con ello la dependencia recíproca, y, 3) da cuenta de la normatividad pre-existente que trasciende a los individuos y sus interacciones, explicando el papel de un agente activo genuinamente empotrado en ella.

Capítulo 3

Enactivismo

3.1. Introducción

En este capítulo me enfoco en analizar el marco enactivo, sus aportaciones específicas a la cognición socio-moral y sus implicaciones normativas. Este análisis lo elaboro a partir de las implicaciones que involucra mi marco integrativo, en donde la parte enactiva emplea una metodología *relacional*, su entidad núcleo es la *acción* y sus principales niveles de análisis son lo *perceptual* y lo *supra-personal*.

La presentación de temas es de la siguiente manera: primero, expongo los planteamientos evolucionistas asociados con la escuela de la “construcción de nicho”, en donde destaco la contribución de elementos corporizados que, en sincronía con procesos culturales, favorecieron decisivamente la evolución de la cognición socio-moral. Segundo, presento el modelo cognitivo del enactivismo, lo que permitirá ver las aportaciones de concebir la cognición como interacciones agente-entorno propiciadas por bucles de percepción-acción. Tercero, respecto a la cognición social, desarrollo la explicación de la “interacción socio-cognitiva”, de lo que derivan los principales requerimientos de la normatividad moral enactiva. Con ello, doy paso a la cuarta sección enfocada en la “cognición moral enactiva”, en donde abordo puntualmente dos temas: 1) la propuesta de una normatividad moral enactiva que plantea un principio moral básico, a saber, *el cuidado de las autonomías*, lo que resalta la importancia tanto de las autonomías individuales como de las heteronomías sociales, y; 2) la evidencia empírica que apuntala las intervenciones corporizadas en la cognición socio-moral humana, principalmente la evidencia pertinente a la perspectiva de 2ª persona, al involucramiento emocional, la intencionalidad compartida y a la relevancia de la interacción.

Si mis consideraciones son correctas, de ello planteo claramente las aportaciones del marco enactivo al tema de interés, lo que me permitirá responder a la pregunta específica: ¿qué resuelve empíricamente el enactivismo que el computacionalismo no? al menos, según mi propuesta. Con esto en mente, paso a la primera sección de este capítulo llamada “Evolución”.

3.2. Evolución

En esta sección, mi intención es relacionar la emergencia y consolidación de las capacidades asociadas al desarrollo moral con la hipótesis de la construcción de nicho. Para ello, en primer lugar, expongo los planteamientos generales de dicha hipótesis, resaltando la construcción del “nicho cultural” propiamente humano que supone un aprendizaje intenso y continuo, el cual es de un tipo específico: el *aprendizaje cultural* o *aprendizaje social*. En segundo lugar, se verá que este tipo de aprendizaje fue posible al menos por tres elementos: 1) la comunicación (simbólica), 2) la intersubjetividad y, 3) el lenguaje, por tanto, es preciso revisar la relación prevaleciente entre ellos, y, finalmente, considerar su vínculo con la evolución moral. En ello, se verá que las prácticas de aloparentalidad jugaron un papel crucial, pues junto con estos elementos, se configuró el escenario que impulsó la emergencia de capacidades socio-cognitivas y de conductas prosociales humanas.

Así, el argumento es que la moralidad, como parte del nicho cultural humano, se transmite mediante la herencia cultural habilitada por el aprendizaje social, y este a su vez es facilitado por la comunicación (simbólica) y la intersubjetividad; características favorecidas por la aloparentalidad, la imitación y la multimodalidad. De manera relevante, se verá que el origen de este nicho dinámico es la *simpatía*, por lo que es plausible pensar que el núcleo biológico de rasgos heredables se encuentra en el origen de ello.

Siguiendo esta línea, se plantea que el nicho cultural humano resulta determinante en la formación y consolidación de la cognición humana y, en particular, de la cognición socio-moral, cuyo desarrollo supone dicho aprendizaje cultural habilitado por interacciones constantes y maleables con el medio social. Por esta razón, se plantea que la evolución moral dependió crucialmente de la comunicación, de la intersubjetividad y del lenguaje, en tanto que es un fenómeno intensamente cultural.

Previo a continuar, veamos brevemente la relación entre el enactivismo y la hipótesis de la construcción de nicho.

3.2.1. Enactivismo y construcción de nicho

Primero, una restricción debe ser explicitada. En principio, la línea argumentativa del enactivismo habla de una adaptación *in situ* y no a la carga adaptativa evolutiva de los

enfoques tradicionales. Pero, dado que mi propuesta otorga un papel determinante al nivel supra-personal, la herencia ecológica (natural y social) debe ser explicada.

En este contexto, es preciso buscar una hipótesis evolucionista que se ajuste potencialmente a los principios enactivos, lo que significa que requerimos una teoría de la evolución que justifique la co-constitución agente-entorno como un tipo de estrategia adaptativa que explota y aprovecha al ambiente en el curso de la evolución de la especie.

La hipótesis de la construcción de nicho cumple con estos requerimientos, pues reconoce tal co-constitución agente-entorno, en la que los agentes participan de manera *activa y relacional* con su entorno, remodelándolo y estructurándolo para las generaciones posteriores. Así, la construcción de nicho pone en primer plano la fuerza evolutiva asociada con los procesos culturales, pero, en origen, está anclada a la evolución biológica, lo que permite proponer la co-evolución bio-cultural del conocimiento moral que planteo.

Con esto en mente, ahora reviso los principios de la hipótesis de la construcción de nicho.

3.2.2. Construcción de Nicho

La escuela de la construcción de nicho destaca por relacionar potencialmente la evolución con la influencia cultural, dando así un papel preponderante e irrevocable a la cultura en sus explicaciones evolucionistas. En consecuencia, reconoce el carácter *activo* de la evolución humana, con lo que postula que el humano se adapta a diversos entornos debido a sus altas capacidades sociales y culturales y de esta manera otorga un papel crucial al *aprendizaje cultural*.

En contraste con la visión de la teoría evolutiva clásica, en la que los organismos se adaptan (o no) de forma *pasiva* a un entorno cambiante, la construcción de nicho coloca a los organismos en el centro de la *acción*, construyendo o modificando activamente su entorno con lo que alteran la dinámica evolutiva (Sanso-Guerrero, 2018, p.2).

Así, la construcción de nicho se entiende como un tipo de ingeniería epistémica que los grupos de organismos despliegan con el fin de conducirse a nuevos ciclos de retroalimentación orientados al aprovechamiento de las posibilidades del medio, o bien, de los recursos disponibles que se hallan en él. Las conductas efectuadas por los organismos en estos procesos deben ser consideradas como conductas *adaptativas*, puesto que la explicación de la construcción de nicho está íntegramente anclada a la teoría de la evolución

biológica, sin embargo, a su vez, debe reconocerse el papel activo y la co-constitución con procesos culturales evolutivos.

En este punto, es preciso aclarar el concepto de “conductas *adaptativas*”, pues ha prevalecido una dialéctica entre la selección natural y la construcción de nicho que requiere de una reformulación del concepto de *adaptación* que no se limite a la noción de una causalidad unidireccional en que los ambientes filtran los genes de ascendientes a descendientes (ver Álvarez, 2013, p.351):

“Con los principios conjugados de selección natural y construcción de nicho es necesario revisar el concepto de *adaptación*, que con la sola referencia a la causalidad de la selección quedaba reducida a la influencia del medio sobre los organismos. La construcción de nichos obliga a reconocer *el carácter dual* de la adaptación. De esto también existen antecedentes, como de casi todo. El gran psicólogo evolutivo del siglo pasado, Jean Piaget, que empezó siendo malacólogo, comprendió la adaptación como la *síntesis* de dos procesos opuestos, a saber, la *acomodación* del organismo al medio y la *asimilación* del medio por el organismo (Cf J. Piaget 1969). La semejanza entre *acomodación* y selección natural, y entre *asimilación* y construcción de nicho, no es un paralelo perfecto, pero no es posible pasarla por alto (Cf. S.T. Parker 2005)”. (Álvarez, 2013, p. 352, las cursivas son mías).

En particular, la *acomodación* del organismo al entorno es claramente asociable a los principios de la selección natural, pero la *asimilación* del entorno por el organismo es asociable a “un principio técnico según el cual el medio o ambiente es también una fuente de recursos transformables y al principio de construcción de nicho. La interacción no trivial entre selección natural y construcción de nicho con las consecuentes herencias genética y ambiental constituye una dinámica interna a un proceso de *adaptación* que ya no se reduce a la *acomodación*” (Álvarez, 2013, p.352, las cursivas son mías).

El planteamiento apunta a reconocer la interacción no trivial entre la selección natural y la construcción de nicho que deriva en dos herencias, la genética y la ambiental, lo que constituye una dinámica de un proceso *dual* de *adaptación* que ya no se reduce a la causalidad unidireccional de selección, en donde sólo se reconoce la influencia ejercida del entorno sobre los organismos. En cambio, debe reconocerse la fuerte influencia recíproca prevaleciente entre el organismo y el entorno, lo que da pie a una genuina co-constitución entre ellos.

De esta manera, la tesis de la construcción de nicho refiere a este proceso co-constitutivo en la naturaleza, donde el organismo modifica los componentes vivos y no vivos del medio

ambiente a través de su propia actividad, esto es, mediante sistemas de explotación/aprovechamiento contextualizados ontogénicamente el organismo altera su ambiente, lo que provoca que su propio desarrollo se vea afectado por tales modificaciones ambientales, lo cual, permite establecer el papel que el ambiente juega en este proceso de retroalimentación así como la participación que el organismo tiene en el mismo. Respecto al “nicho humano”, el hombre modifica su medio ambiente principalmente a través de procesos ontogenéticos, sociales y culturales.

Ahora bien, la modificación ambiental que surge de la construcción de nichos causa cambios evolutivos en el organismo, los que se transmiten como legados a través de generaciones (herencia ecológica). No obstante, tal herencia de nichos requiere de la capacidad de los organismos para responder a los mismos. Como vemos, los organismos, en la dinámica perpetua del proceso evolutivo, afectan al ambiente y, a su vez, estas afectaciones ambientales influyen en el proceso evolutivo de los organismos. Así, pues, la construcción de nichos se da porque las especies buscan maximizar el uso de los recursos y, en consecuencia, la probabilidad de supervivencia de una progenie se proyecta en un futuro distante.

En el caso de los agentes humanos, esto sería la construcción del “nicho cultural” y una fuente fundamental de su herencia es la comunicación, puesto que ésta hace posible el *aprendizaje cultural*. Es decir, dado que el proceso de evolución de un nicho cultural no se limita a la herencia genética, sino que también incluye la herencia ontogenética, conductual y simbólica, y, debido a que las actividades culturales del ser humano han influido en la evolución genética, entonces, la construcción de nicho potencializa los efectos de la cultura para la evolución humana⁵¹ (Flynn, Laland, Kendal y Kendal, 2013).

De esta manera, se genera un nicho cognitivo en el que habitan y se desarrollan los seres humanos; este nicho cognitivo es un modo de supervivencia caracterizado por la manipulación del ambiente mediante el razonamiento y la cooperación social, siendo sus elementos principales el lenguaje, la dependencia infantil y la inversión parental. Más adelante presento los argumentos que sugieren que estos elementos emergieron de practicar la *aloparentalidad*.

⁵¹ Desde las explicaciones de la *herencia dual*, un ejemplo paradigma es la tolerancia a la lactosa en los humanos, que se dio a partir de la aparición de las granjas y de la crianza de las vacas (ver Kiverstein, 2017).

Así, pues, se trata de proponer al organismo como unidad del nicho, como unidad de *acción* conductual que toma decisiones confluyendo como un todo en una dinámica donde el organismo es parte y forma del nicho.

Ahora bien, tras revisar brevemente los principios de la hipótesis de la construcción de nicho, veamos a continuación sus aportaciones directas a la cognición socio-moral.

3.2.3. Aportaciones directas de la hipótesis de la construcción de nicho a la cognición socio-moral humana

Como se ha visto, se trata de hablar de dos tipos de herencias, la genética que incluye los rasgos biológicos heredables y la ecológica que refiere a la herencia de las modificaciones en el medio. En esta sección me enfoco en la segunda; en la herencia ecológica.

Para ello, es preciso subrayar que la herencia ecológica implica la herencia cultural, la cual es posible gracias al aprendizaje social y los elementos que le habilitan son la comunicación y la intersubjetividad, mismas que fueron favorecidas fundamentalmente por la imitación⁵². Pero ¿qué impulsó específicamente la evolución de la comunicación y de la intersubjetividad humana? se verá que la respuesta es la crianza cooperativa o *aloparentalidad*, la cual encuentra su origen en la *simpatía*.

Esto en conjunto develó un nicho dinámico que en su continua interacción definió el escenario para la emergencia de las conductas prosociales y cooperativas humanas; proceso que involucró los rasgos genéticamente heredables (instintos sociales, filiales, parentales, empatía y los mecanismos encargados de nuestras orientaciones normativas generales) y a su vez, la herencia ecológica transmitida mediante el aprendizaje social.

Así, para mis propósitos, tomo esto como base para dar cuenta de los elementos que pudieron dar lugar al conocimiento moral general dotado evolutivamente, al menos de manera parcial.

⁵² El primero en reconocer el papel tan significativo que tiene la *imitación* en el conocimiento humano fue Aristóteles, quien sostenía que el entendimiento comienza mediante la imitación. En sus palabras: “la propensión natural del hombre, desde la infancia en adelante, a participar en la actividad mimética, lo distingue de otras criaturas, pues él es completamente mimético, y es a través de la mimesis que toma sus primeros pasos hacia la comprensión” (Aristóteles, [1987], p. 34, citado en Etzelmüller, 2016, p. 153).

Para ver esto con detalle, comienzo con los antecedentes evolutivos de la intersubjetividad y de la comunicación simbólica, debido a que jugaron un papel crucial en el desarrollo del *aprendizaje social*.

3.2.4. Antecedente de la intersubjetividad: evolución del uso de símbolos

En general, se considera que la evolución del lenguaje, de la moral y del pensamiento humano se ha basado fundamentalmente en el uso de símbolos, estos entendidos como representaciones socialmente compartidas, convencionales y sistemáticas. Pero, ¿qué impulsó el desarrollo de la comunicación simbólica?

Al respecto, Jordan Zlatev (en Etzelmüller, 2016) ofrece un argumento valioso acerca de la evolución de la comunicación simbólica específica al humano basándose en tres condiciones esencialmente corporizadas: la aloparentalidad, la mimesis corporal y la multimodalidad; condiciones que acomoda de la siguiente manera:

- 1) La crianza cooperativa o *aloparentalidad* se adoptó como una estrategia reproductiva clave que posibilitó la evolución de un grado de *intersubjetividad*.
- 2) La *aloparentalidad* y la *intersubjetividad* propiciaron la emergencia de la conducta pro-social.
- 3) La conducta pro-social se desarrolló en las culturas prelingüísticas a través de la *mímesis corporal* (Donald 1991; Zlatev 2008), que funge como una precondition de la comunicación humana.
- 4) Finalmente, la *multimodalidad* vocal-manual fue esencial para la transición de un modo de comunicación predominantemente corporal e icónica a un modo predominantemente (aunque no exclusivamente) vocal y simbólico.

La propuesta implica una extensa co-evolución bio-cultural y se basa en modelos darwinianos extendidos, como la “teoría de la selección multinivel”, que opera tanto en individuos como en grupos, con compensaciones complejas entre sí; características consistentes con la hipótesis de la construcción de nicho. Revisemos con detalle la propuesta.

Como punto de partida, Zlatev define el uso de símbolos como la producción y comprensión de signos que tienen las siguientes características: a) son *representativos*, *i.e.* son entendidos por sus usuarios como denotando objetos, propiedades y eventos generales o específicos (Sonesson 2007); b) son *convencionales*, es decir, son compartidos dentro de

una comunidad (Itkonen 2003), aunque no necesariamente “arbitrarios”; c) son expresados junto con la *intención comunicativa* (Tomasello 1999; Zlatev 2013) y; d) son *sistemáticamente* interrelacionados (Saussure [1916] 1983; Diácono 1997).

La propuesta de Zlatev depende de dos condiciones: 1) la *mímesis corporal*, que es condición clave, aunque no suficiente, para la explicación evolutiva de la comunicación simbólica humana, y, 2) la *selección multinivel*, la cual revive y elabora la idea darwiniana respecto a que la selección puede operar en grupos completos, y no solo en individuos y sus genes (Richerson y Boyd 2005; Sober y Wilson 1999).

En particular, la hipótesis de la *mímesis corporal* establece que “una adaptación para mejorar el control volitivo del cuerpo les dio a nuestros antepasados ventajas en los dominios de la imitación, la empatía y la comunicación intencional (gestual). Esto allanó el camino para la evolución del lenguaje, sin ninguna otra adaptación biológica. . . además de un mejor control vocal” (Zlatev 2014, p.166, citado en Etzelmüller, 2016, p. 153).

Tomando esto en cuenta, en adelante, se defiende una potencial liga entre la crianza cooperativa (aloparentalidad) y el surgimiento de las capacidades socio-morales específicas al humano. Concretamente, se plantea que las prácticas de aloparentalidad configuraron el escenario que impulsó la emergencia de dichas capacidades.

El punto de partida es preguntar ¿qué fue primero, el lenguaje o la intersubjetividad?

En contraste con algunas teorías que formulan la intersubjetividad como una especie de subproducto del lenguaje⁵³, Zlatev sostiene que la intersubjetividad precede al lenguaje. Este planteamiento encuentra apoyo en las explicaciones que sostienen que la vida en grupo permitió la auto-domesticación; rasgo específicamente humano que derivó y contribuyó en la evolución de una cascada de características más sofisticadas, entre ellas, el lenguaje y la moral (Chen y Moyzis, 2018).

Ahora bien, Zlatev, valiéndose de la teoría de la selección multinivel, recurre a la propuesta de “aloparentalidad” de Hrdy (2009), quien propone que lo que inició la cascada de procesos que condujo a un mayor tamaño del cerebro, a conductas cooperativas y prosociales, a habilidades cognitivas específicas de los humanos y, en última instancia, al lenguaje, fue

⁵³ Por ejemplo, Deacon (1997) con su hipótesis del “contrato sexual” y Dunbar (1996) con su teoría del “acicalamiento vocal y chisme” (en Etzelmüller, 2016, p.159).

un cambio en la estrategia reproductiva y de crianza: la *aloparentalidad* o crianza cooperativa.

En particular, Hrdy sostiene que la *crianza cooperativa* fue la condición preexistente que permitió la evolución de los rasgos cognitivos humanos y del lenguaje. La hipótesis es que las tendencias de crianza de las madres se extendieron a otros miembros del grupo, incluidos los padres, entre otros, lo que permitió que los recursos, el alimento y el cuidado de los niños se distribuyeran de manera más equitativa dentro del grupo.

Esto sugiere un escenario co-evolutivo de infantes más inmaduros, infancias prolongadas y, por tanto, más necesidad de cuidado y de provisión compartida, en el que la aloparentalidad no fue el único factor inicial, como sugirió Hrdy, sino que también fue facilitado por una mayor altricialidad (es decir, bebés muy inmaduros). (Etzelmüller, 2016, p.162).

Hrdy basa su teoría en las prácticas de cuidado infantil de los grupos de cazadores-recolectores, a quienes concibe como los modelos más cercanos de lo que pudo haber sido la vida de nuestros antepasados pre-agrícolas: “Es significativo que en sociedades cazadoras-recolectoras muy distintas cultural, geográfica y ambientalmente como Aka, Efé, Kung San (África Central), Himba (África Occidental), Yanomamo (Venezuela) y Agta (Filipinas), el cuidado se comparte entre madres y aloparentales” (*Ibid.*, p.163).

En particular, la aloparentalidad parece promover la preocupación y el cuidado por los demás, lo que deriva en favorecer la intersubjetividad humana y en propiciar la conducta prosocial. De esta manera, la teoría de Hrdy busca ajustarse a una explicación consistente de cómo la intersubjetividad humana pudo emerger mediante la crianza cooperativa que supone una necesidad evolutiva.

El planteamiento señala que en el contexto de la reproducción comunitaria y el cuidado compartido de los infantes se sitúa la *simpatía* directa con las expresiones reales de los sentimientos de los otros, lo que implicó un "proceso evolutivo de auto-reforzamiento de padres y alopadres que son más sensibles a las señales infantiles y bebés que son mejores para emitirlos" (Hrdy 2009, p. 220), constituyendo así un proceso de adaptación que afectó a todo el nicho social de interacción.

Según Hrdy, al crecer en el contexto de la aloparentalidad, las capacidades del infante humano fueron seleccionadas y se desarrollaron con una mayor comprensión de la toma

de perspectiva y de la autoconciencia, así como también, de sus aspectos motivacionales cargados de emociones y fundamentados en expresiones básicas de la corporización humana, lo que benefició decisivamente las capacidades socio-cognitivas.

De manera relevante, en la psicología del desarrollo, la propuesta de la “intersubjetividad” de Trevarthen (1979) captura directamente este planteamiento al sostener que: “nacemos para generar estados cambiantes de autoconciencia, para mostrárselos a otras personas, y provocar en ellos interés y respuestas afectivas” (Trevarthen, 2011, p. 119).

Así, la hiper-socialidad humana se caracterizó parcialmente por la cooperación y por la conducta prosocial, lo que encontró sus precedentes socio-cognitivos en las prácticas alopARENTALES, lo que puede explicarse de la siguiente manera:

Trevarthen y Bråten distinguen, esquemáticamente, entre (1) *intersubjetividad primaria*, que surge desde los primeros meses de vida en adelante y se basa en la “simpatía directa con las expresiones reales de los sentimientos de otros en contacto íntimo recíproco sujeto-sujeto”; (2) *intersubjetividad secundaria*, que se da a partir de los 9 meses de edad e involucra “objetos de atención conjunta y referencia emocional, que se ponen en juego dentro de relaciones confiadas de compañerismo . . . a veces invitando al aprendizaje imitativo”; y, (3) *intersubjetividad terciaria*, basada en “conversaciones simbólicas con compañeros reales o virtuales” (Bråten, 2007, p. 3).

Zlatev asocia este planteamiento con las etapas evolutivas de la siguiente manera: (1') Los rasgos de intersubjetividad específicos de los humanos evolucionaron primero en el nivel de "simpatía directa" entre el niño y los alopAdres; (2') posteriormente se extendieron a las habilidades cognitivas de intencionalidad compartida y mimesis (diádica y triádica), y; (3') sólo (mucho) más tarde, con la evolución del protolenguaje y del lenguaje, fueron las “conversaciones simbólicas” y el tipo de normas morales (Zlatev en Etzelmüller, 2016, p.165).

Así, la hipótesis de la alopARENTALIDAD es compatible con las explicaciones de las conductas prosociales y con las explicaciones pertinentes al desarrollo social de los infantes y, en esta línea, también presta apoyo a las explicaciones del desarrollo de las capacidades para establecer escenarios compartidos con los otros.

De ello, el argumento puntual es que la crianza cooperativa impulsó la transformación de capacidades comportamentales más básicas del linaje homínido⁵⁴ a capacidades más complejas orientadas a fines cooperativos y a la emergencia de la prosocialidad; como conductas de ayuda, aversión a la inequidad, comunicación intencional compartida y a la cultura acumulativa en general (Burkart et al., 2009). Esto permitió que cada una de estas habilidades se potencializara socialmente posibilitando la creación de escenarios de actividad conjunta y evolución cultural compleja, estableciendo en lo definicional el marco de la intencionalidad compartida que *gatilla* la subjetividad, la intersubjetividad e incluso la postura ética del ser humano ante los demás (Segovia-Cuellar, 2017, p.13).

Así, pues, esta descripción plantea una potencial liga entre la crianza cooperativa y el surgimiento de las capacidades socio-morales específicas al humano, debido a que, desde este punto de vista, la aloparentalidad (junto con la intersubjetividad derivada de ello) proporcionó las bases para la emergencia de las conductas cooperativas y prosociales.

Ahora bien, el segundo paso en mi tratamiento respecto a la evolución de las capacidades socio-morales humanas asociadas con los principios de la construcción de nicho, es detallar cómo se relaciona lo visto hasta aquí con el lenguaje, esto debido al papel tan preponderante que posee el *aprendizaje social* en la herencia cultural, el cual se potencializó a partir de la comunicación, por lo que es necesario revisar los elementos que dieron lugar a ello.

En particular, se verá que la aloparentalidad y la intersubjetividad humana, en conjunto, se beneficiaron de la *imitación* entre agentes, lo que a su vez propició la comunicación y en última instancia, el lenguaje. Por tanto, ahora corresponde ver el papel de la *mímesis corporal* en el lenguaje, lo que aborda el último punto de la propuesta de Zlatev: la *multimodalidad*.

3.2.5. Transición de la mímesis corporal al lenguaje

⁵⁴ Capacidades derivadas de su libertad motora y repertorio de acciones intencionales, lo que incluye la posibilidad del aprendizaje social, la utilización de herramientas sencillas, el entendimiento perceptual básico de las intenciones y objetivos del comportamiento en otros y la comunicación gestual (Donald, 1991; Leroi-Gourhan, 1993; Zlatev et al., 2005; Malafouris, 2013).

Primero, el concepto de la *mímesis corporal* es rico en relaciones y extensiones, pues es capaz de contribuir, a partir de la habilidad motora, a la cognición social, e incluso a la cultura, puesto que propicia el aprendizaje social.

En particular, la *mímesis corporal* contribuye directamente a dos de las cuatro características del uso de símbolos que propone el autor, concretamente a: 1) su carácter representativo y, 2) a su implicación de intención comunicativa.

Respecto a (1), recordemos que para Zlatev, los símbolos son representativos en el sentido de que denotan objetos, propiedades y eventos generales o específicos. Frente a esta concepción, la *mímesis corporal* contribuye significativamente a ello por ser capaz de crear una relación representacional entre signo y objeto, sobre todo a partir de un fundamento icónico basado en la semejanza.

Respecto a (2), la sugerencia es que los símbolos son expresados junto con la intención comunicativa, la cual es una intención de segundo orden para que el otro capte la intención informativa de primer orden del comunicador. La *mímesis corporal* contribuye a la comprensión de la intención comunicativa en el sentido de que el sujeto comunicador tiene la intención de que el acto que ejecuta para comunicar represente alguna acción, objeto o evento para un otro, y que el otro reconozca esa intención. Esto contribuye a la atención conjunta, pues, mediante el acto ejecutado del comunicador, ambos ponen atención a un tercer objeto (ej. deseo que veas que veo X), lo que da lugar a la intencionalidad compartida (comunicador y receptor comparten un mismo objetivo). Esta característica es favorecida principalmente por señalamientos, gestos y atención conjunta. A este nivel de *mímesis*, el autor le llama *mímesis tríadica*.

Crucialmente, Zlatev entiende las capacidades de la *mímesis corporal* en capas jerárquicas, es decir, las más altas comprenden a las más básicas, en este caso, (2) comprende a (1) y, en esta línea, argumenta que la combinación de estas dos características es una característica prelingüística específica del ser humano.

Pero, concretamente, ¿cómo pasamos de la *mímesis corporal* a la comunicación simbólica? Y en relación con ello ¿cómo se explican las otras dos características del uso de símbolos, es decir, su carácter convencional y sistemático? Para responder, Zlatev recurre a la tesis de la *evolución de la comunicación simbólica: una perspectiva corporizada* (Brown, 2012), que justamente aborda el hueco explicativo que existe entre la *mímesis corporal* y la

comunicación simbólica, en donde encuentra al menos tres condiciones relacionadas con la corporización humana: la aloparentalidad, la mimesis corporal y la multimodalidad, las cuales acomoda en una explicación interpretada en tres transiciones (ver Etzelmüller, 2016, p.166):

- *Transición 1*: La transición gradual que hubo entre el modo de expresión meramente corporal (gestual) y el modo de expresión predominantemente vocal puede explicarse desde la propuesta de aloparentalidad de Hrdy que caracteriza el nicho evolutivo en el que la intersubjetividad humana evolucionó.
- *Transición 2*: Este proceso estuvo acompañado de una pérdida parcial de iconicidad y, por lo tanto, dio paso a una mayor confianza en el conocimiento de las *convenciones* lingüísticas.
- *Transición 3*: El autor sugiere que la combinación de estas dos transiciones pudo haber brindado una contribución clave para la evolución de la característica final de los símbolos: las relaciones elaboradas dentro del sistema y su potencial combinatorio (*i.e.*, son sistemáticos).

La razón crucial es que el medio vocal tiene menor potencial para representar el significado mediante la base de similitud, que el medio manual-corpóreo (Brown, 2012; Fay, Arbib y Garrod 2013). Esta característica significa que el medio vocal es capaz de bajar la probabilidad de *arbitrariedad* en los códigos, pues, en cierta medida, evita las ambigüedades que suelen tener los códigos icónicos favorecidos por el medio manual-corpóreo (que basa el significado en la similitud). Además, el medio vocal posee el potencial de ampliar el repertorio de mensajes a comunicar que ya incluían los códigos meramente icónicos.

Así, la *transición 1* puede entenderse desde la aloparentalidad, que se centra en el nicho evolutivo de la intersubjetividad humana favorecida por la imitación. Como se vio, sus etapas evolutivas comenzaron en el nivel de la *simpatía* directa que emerge entre infantes y cuidadores (alopadres), posteriormente dicha simpatía se extendió a capacidades cognitivas intencionales y compartidas y, mucho más tarde, estas habilidades evolucionaron, junto con el proto-lenguaje y el lenguaje, a las conversaciones simbólicas, e incluso, alcanzaron el nivel de cognición requerido para albergar normas del tipo moral. La idea central es que debido a que la crianza cooperativa catapultó la transformación de capacidades comportamentales más básicas a capacidades más sofisticadas, tales como

la intersubjetividad propiamente humana, las conductas cooperativas y prosociales, la ayuda, la comunicación intencional compartida y la cultura acumulativa, entre otras, fue capaz de potencializar un marco complejo de evolución cultural que comprende desde las habilidades motoras más básicas hasta las capacidades simbólicas propiamente humanas.

En este contexto, la *transición 2*, que implica la *convencionalidad* del uso de símbolos, se explica mediante la mimesis corporal que dio lugar a lo que Donald (1991) llama la “cultura mimética” basada en la producción de herramientas, la pedagogía, la división del trabajo, la caza y la recolección cooperativa. Estas capacidades se beneficiaron de una pérdida parcial de la iconicidad, lo que dio paso a una plétora de *convenciones* simbólicas en los códigos. Todo ello presupone una comunicación mimética, multimodal pero predominantemente gestual, con intenciones comunicativas: “mimesis triádica” (es importante destacar que esto aún no era un lenguaje, y posiblemente ni siquiera un proto-lenguaje).

El tipo de multimodalidad que supone la mimesis corporal se beneficia tanto del canal corporal-visual como del canal vocal-auditivo, ambos trabajando en coordinación, pero a su vez, operando como dos modos de expresión complementarios. Así, su naturaleza es de un tipo interactivo capaz de ampliar las capacidades comunicativas.

Ahora bien, la *transición 3* respecto al cambio gradual y prolongado del proto-lenguaje gestual al proto-lenguaje vocal, fue favorecido por la naturaleza multimodal de la mimesis corporal y las diferentes capacidades intrínsecas del gesto y del habla para la iconicidad. Inicialmente, el gesto y la vocalización pudieron haber tenido funciones diferentes (representación y afiliación, respectivamente), pero con el tiempo, las vocalizaciones habrían sido “reinterpretadas” como signos comunicativos, dada la estrecha sincronización de la comunicación multimodal, gestual y vocal, dando lugar así a la *sistematicidad*:

Con los grupos sociales más grandes, más organizados y tecnológicamente avanzados del Homo sapiens anatómicamente modernos de hace 200 mil años, un vocabulario más amplio de expresiones convencionales, que dependía cada vez más del canal vocal, habría proporcionado el ímpetu para la sistematización. Así habría surgido la capacidad para aprender y usar relaciones intrincadas signo-signo: la característica final de la definición del uso de símbolos utilizada por Zlatev [la sistematización] (Etzelmüller, 2016, p. 169).

El planteamiento brinda apoyo a la característica de *sistematización* del tratamiento de Zlatev y apunta a considerar que el potencial humano para la construcción de sistemas puede ser en cierto sentido “innato” y, sin embargo, como sostiene Arbib (2005), ser un

potencial pos-biológico que emerge mediante procesos de la evolución cultural. Puntualmente, la evolución cultural contribuyó decisivamente a la *sistematización* de los símbolos, debido a que la transición del modo de expresión gestual al modo de expresión vocal y el surgimiento de las convenciones simbólicas, fueron factores que en su evolución exigieron la emergencia de relaciones elaboradas signo-signo, así como también su potencial combinatorio dentro de los códigos. De esta manera, se sugiere que las *transiciones 1 y 2* contribuyeron crucialmente a la sistematización de la comunicación simbólica.

Así, pues, si mis interpretaciones son correctas, hasta aquí se han cumplido las cuatro características que dan pie a la comunicación simbólica, según Zlatev, en donde el uso de símbolos se define como la producción y comprensión de signos que son *representativos, convencionales*, expresados junto con la *intención comunicativa* y son *sistemáticamente* interrelacionados, lo que nos acerca a una posible explicación de la transición que ocurrió entre la comunicación gestual y la comunicación vocal.

En ello, se vio que la comunicación simbólica y la imitación (y las condiciones corporizadas que implican: aloparentalidad, mimesis corporal y multimodalidad), fueron factores que favorecieron decisivamente las prácticas intersubjetivas, las cuales a su vez impulsaron el aprendizaje social, que es el factor clave para la herencia cultural.

Esto constituiría un nicho dinámico en el que tuvo lugar la emergencia de las capacidades socio-cognitivas y conductas prosociales humanas, lo que estaría crucialmente enraizado en el núcleo biológico de rasgos heredables compuesto por la empatía, por instintos sociales, parentales, filiales y por aquellos mecanismos que subyacen a lo que conforma nuestras orientaciones normativas generales.

En este sentido, entonces, estamos hablando de “una red de interacción de múltiples componentes (i.e., genes, productos, sustratos, elementos celulares, extracelulares, ambientales), que permanece estable en el tiempo, pero sin duda en constante desarrollo, pues se reconstruye generación tras generación, estableciendo así sistemas de desarrollo diferenciados. [Aquí] se reclama una democracia causal, que permita identificar las formas vivas como sistemas en desarrollo que recapitulan una y otra vez, un proceso de interacción en red que incluye elementos de todo tipo durante la formación de los sistemas biológicos” (Segovia-Cuellar, 2017, pp.6-7).

Esto permite aproximarnos a una descripción (de *grano grueso*) de las condiciones evolutivas del nicho cultural que propició la intersubjetividad humana y con ello las conductas cooperativas, prosociales, de ayuda, la aversión a la inequidad, etc., elementos que *gatillan* el aprendizaje social y potencializan la moral propiamente humana.

Así, pues, si mis consideraciones son correctas es plausible pensar en la herencia genética y la herencia ecológica co-evolucionando en un marco de interacción capaz de dar pie a un tipo de conocimiento moral general dotado evolutivamente, conformado por aquellas orientaciones normativas generales.

3.2.6. Consecuencias

(A manera de recapitulación, recordemos que en el capítulo 2 se argumentó que la herencia genética nos dota con un núcleo biológico de rasgos heredables que incluye empatía e instintos sociales integrados con los mecanismos que subyacen a lo que conforma nuestras orientaciones normativas generales provenientes de las estrategias exitosas y/o perniciosas de nuestros antepasados).

1. La herencia ecológica cultural depende del aprendizaje social, y los factores que le posibilitaron son la comunicación simbólica y la intersubjetividad.
2. A su vez, estos se vieron favorecidos fundamentalmente por la imitación.
3. La aloparentalidad contribuyó significativamente a configurar este nicho dinámico que en conjunto dio lugar a la emergencia de las capacidades socio-morales humanas; en donde las capacidades sociales fueron potenciadas por la empatía directa con la expresión de los sentimientos de otros y las capacidades morales serían esencialmente las conductas cooperativas y prosociales.

Ahora, corresponde ver los supuestos cognitivos que habilitan el modelo de la cognición enactiva.

3.3. Cognición enactiva: modelo cognitivo

El enactivismo parte de una metodología relacional que entiende la cognición como las interacciones dinámicas y significativas que se dan entre el agente y su entorno, o bien, como *procesos de hacer-sentido (sense-making)* habilitados por acoplamientos sensorio-motores agente-ambiente.

De ello se sigue que el modelo cognitivo del enactivismo es un modelo (no lineal) que funciona a manera de bucles continuos de:

percepción \leftrightarrow acción

En donde la percepción y la acción están intrínsecamente unidas, es decir, la percepción y la acción no se conciben como modalmente separadas, lo que tiene implicaciones cruciales en la cognición socio-moral. Para comprender su modelo cognitivo (a manera de bucle percepción \leftrightarrow acción) es necesario considerar los principios teóricos que se vieron detalladamente en la sección de “Antecedentes”.

Ahora bien, el marco enactivo otorga *autonomía* a los agentes y al nivel supra-personal (interacciones, entorno), por lo que este concepto debe ser definido y no confundirse con la concepción tradicional de la autonomía kantiana.

Primero, el concepto de *autonomía* enactiva posee dos aspectos constituyentes: el *individual* y el *relacional*. El aspecto *individual* refiere a la red de procesos auto-constitutivos que configura la propia identidad del organismo, *i.e.* el agente es autónomo en tanto que se auto-constituye y se auto-mantiene en el tiempo. No obstante, el organismo, a la vez que es *autónomo* es *precario*, característica intrínseca que le exige la perpetua interacción con su entorno para el intercambio de materia y energía que le permita mantener su identidad (o sea, mantenerse vivo).

El aspecto *relacional* surge precisamente de esta continua necesidad de interacción con el entorno, es decir, refiere a la profunda precariedad y vulnerabilidad del organismo que le conduce a una *dependencia constitutiva* con su entorno, puesto que logra preservar su integridad, no sólo por sus procesos auto-constitutivos, sino que también por la continua interacción con su entorno.

Esto es relevante debido a que de esta *dependencia constitutiva* o continua necesidad de interacción con el entorno, emerge una legítima co-constitución agente-entorno o agente-agente, según sea el caso, en donde prevalecen aspectos autónomos, heterónomos y de constitución. Más adelante vuelvo a este tema.

Por el momento, veamos la noción de autonomía enactiva en palabras de van Grunsven (2018):

La autonomía no es lo que sugiere nuestra tradición kantiana, es decir, no es la capacidad de legislar de manera autosuficiente nuestras propias acciones asumiendo un punto de vista racional desapegado que nos separa de nuestros lazos habituales y pre-reflexivos con el mundo. Más bien, los agentes vivos mantienen su autonomía precisamente en virtud de una *dependencia constitutiva* con su entorno. Como tal, la autonomía debe ser reconcebida *relacionalmente* (2018, p. 133, las cursivas son mías).

Bajo estos supuestos, ahora corresponde ver la explicación específica de la cognición social enactiva en la que baso mi propuesta, la cual es la *interacción socio-cognitiva* de Froese y Di Paolo (2011). En particular, retomo esta explicación debido a que otorga las condiciones para la emergencia de una normatividad social, lo que da lugar a la normatividad moral y con ello, a la cognición socio-moral enactiva.

3.3.1. Consecuencias

En conjunto con los principios enactivos que se vieron detalladamente en la sección de Antecedentes, recapitulo las siguientes tres implicaciones de la cognición enactiva:

1. La cognición se efectúa a manera de bucles continuos de percepción \leftrightarrow acción, lo que propicia más posibilidades de actuación y significación para el agente. De manera relevante, no hay separación modal entre la percepción y la acción.
2. La cognición es entendida como procesos de *hacer-sentido* habilitados por las interacciones agente-entorno, de lo que surgen significados y se instauran normatividades.
3. El enactivismo parte de una metodología relacional que suele defender una epistemología interaccionista y su entidad teórica núcleo es la *acción*.

Con esto en mente, veamos la cognición social enactiva.

3.4. Cognición social enactiva

Como primer paso, es preciso recordar los fundamentos generales de la cognición social enactiva:

- 1) La interacción es el recurso de la intersubjetividad.
- 2) Su punto de partida es la perspectiva epistémica de 2ª persona, lo que implica otorgar el rol de *participante* a los interactuantes (y no de meros *observadores*).
- 3) Se beneficia de la *percepción directa* de los estados mentales del otro.

- 4) Se explica mediante el *hacer-sentido participativo* (PSM) que otorga *autonomía* al proceso relacional de la interacción que se da entre dos o más agentes. En consecuencia, el proceso de la interacción habilita y restringe la conducta individual de los participantes; dominio cognitivo relacional no reducible al nivel individual.

Así, el modelo de la cognición-social enactiva es habilitado de la siguiente manera:

En una interacción (digamos, diádica), ambos sujetos se perciben directamente, surge un acoplamiento sensoriomotor entre ellos y actúan en consecuencia (en donde sus acciones son guiadas perceptualmente), lo que resulta en que ambos agentes participan en una mutualidad directa que da lugar a participar en la co-creación de nuevos significados propios de esa interacción específica y por tanto ambos responden a exigencias concretas de la situación *on-line*. Con ello, ambos interactuantes participan y responden a la mutualidad que la interacción en sí misma exige.

De esta manera, tanto el nivel de la interacción como la mirada de normatividades y significados compartidos que surgen de él, son un dominio cognitivo irreductible al nivel individual. Con esto en mente, ahora explico la especificidad de la interacción socio-cognitiva.

3.4.1. Interacción socio-cognitiva

Expongo la interacción socio-cognitiva con dos objetivos: 1) establecer las condiciones necesarias de la normatividad social enactiva, lo que resulta ser la base para la emergencia de la normatividad moral y, 2) para explicar los elementos cognitivos implicados en la cognición socio-moral enactiva. En adición, la explicación de la interacción socio-cognitiva contribuye a entender cómo se le otorga autonomía al nivel supra-personal, en este caso, a la interacción social.

Concretamente, la interacción socio-cognitiva es propuesta por Froese y Di Paolo en 2011 y es una versión enriquecida del *hacer-sentido participativo* (PSM). Se considera una versión enriquecida debido a que suma a los principios del PSM la exigencia de que los agentes deben “percibir al *otro* como agente autónomo” en la interacción. Veamos esta característica con más detalle.

Froese y Di Paolo ofrecen una revisión crítica del enfoque enactivo de la interacción social y proponen una forma específica de agencia capaz de constituir un dominio cognitivo social.

Para ello, los autores establecen dos condiciones necesarias que los interactuantes deben cumplir para que una interacción sea considerada una interacción social en toda regla:

- 1) que los agentes sean capaces de generar valores no-metabólicos y relacionados a *otros* agentes, lo que da lugar a
- 2) que los agentes sean capaces de *percibir al otro como agente autónomo*, i.e. con su propio punto de vista, con sus propias intenciones, etc.

Estas dos condiciones implican crucialmente que el agente que interactúa debe *hacer-sentido* del otro agente como tal, i.e. percibirlo como *otro* agente autónomo en su propio derecho y, para ello, el agente debe ser capaz de generar valores que vayan más allá de su realización y regeneración física y en su lugar crear valores genuinamente relacionados a *otros* agentes.

Tal característica apunta a una des-centralización de significado, pues ahora el agente es capaz de lograr hacer-sentido de otros de una manera que no se confina a las necesidades metabólicas y auto-dirigidas, y con ello es capaz de percibir al otro como *otro*, lo que es regido por una normatividad relacionada a *otros*.

En esta línea, los valores que gobiernan el desarrollo de las interacciones sociales preservan una relativa independencia con respecto a las normas de realización y regeneración física del organismo. Así, la normatividad social debe ser una normatividad no determinada por valores metabólicos.

Para comprender tal normatividad, es preciso explicar las tres capas que el enactivismo reconoce como normativas, mismas que mantienen relaciones de interdependencia entre sí, pero a su vez, presentan diversos desacoplamientos entre ellas: *normatividad biológica*, *normatividad sensorio-motora* y *normatividad social*.

3.4.2. Normatividades enactivas

Primero, la normatividad surge de las interacciones agente-entorno en las que el agente *hace-sentido* de su entorno en virtud de su propia constitución y de sus necesidades. En estas interacciones el entorno se vuelve un mundo irreductiblemente significativo para el agente debido a que percibe sus características como oportunidades de acción, interacción y respuesta, de tal manera que el entorno es un medio de posibilidades cargadas de valor en tanto que tales posibilidades pueden ser buenas, malas o neutras para el organismo, pues mediante ellas puede renovar (o no) los procesos que constituyen su propia identidad,

sea física o social. Así, ser un ser autónomo *hacedor-de-sentido* es ser una fuente de *hacer-valores*: "algo adquiere significado para un organismo en la medida en que se relaciona (positiva o negativamente) con la normatividad que mantiene la integridad del organismo" (Thompson, 2007, p. 70).

En este contexto, el enactivismo entiende la normatividad dependiendo de la constitución de los agentes; entre más compleja es la constitución de los agentes más compleja es la normatividad de la que es capaz el organismo. Con estas características en mente, a continuación, explico las distintas normatividades enactivas.

-Normatividad biológica (normatividad individual). Este nivel de normatividad es determinado por procesos metabólicos y de regeneración física, por tanto, refiere exclusivamente a valores metabólicos y dirigidos al organismo mismo. Esto debido a que es una normatividad instaurada por procesos de *hacer-sentido* y de adaptividad dedicados al auto-mantenimiento del agente. Entonces, en este nivel, un ser vivo como sistema autónomo es un sistema auto-determinado en lugar de ser un sistema determinado por el exterior. Crucialmente, aquí los objetivos regulatorios son determinados por las necesidades metabólicas, de tal manera que este nivel de normatividad es en *pos* de mantener vivo al organismo (p. ej. ver Froese y Di Paolo, 2011).

-Normatividad sensoriomotora (interacciones sensorio-motoras). Este nivel de normatividad se fundamenta en sistemas relativamente desacoplados de procesos metabólicos, por lo que ya no son determinados exclusivamente por valores metabólicos, por ejemplo, el sistema inmunológico o el sistema nervioso. Aquí, el foco es el sistema nervioso por ser el que gobierna las interacciones sensoriomotoras, pues es el que permite al agente crear bucles de percepción-acción mediante los cuales propicia más posibilidades de actuación y significación. En este nivel de normatividad podemos comenzar a hablar de valores no-metabólicos, y por ello, de valores no-dirigidos al mero auto-mantenimiento del agente, lo que significa que a partir de este nivel podemos empezar a des-centrar la noción de valores dirigidos-a-uno para arribar a valores dirigidos-a-otro. A continuación, lo explico con más detalle.

Debido a la relativa independencia del sistema nervioso respecto a los procesos constructivos metabólicos, la regulación normativa de la interacción sensoriomotora no es determinada por necesidades energéticas y materiales básicas. De esta manera, el sistema nervioso permite la aparición de nuevas dinámicas autónomas, de nuevos patrones de

percepción y acción que están relativamente desacoplados de los procesos metabólicos, de tal modo que la regulación del comportamiento sensoriomotor está libre de los estrictos límites de la normatividad básica dirigida-a-uno y puede ser en cambio relacionada a algo diferente; a otro agente. El quid del argumento es que la normatividad sensoriomotora comienza a desvincularse de la creación de valores metabólicos dirigidos a la regeneración y mantenimiento del agente mismo, dando paso a la creación de valores dirigidos a *otros* con lo que adquieren patrones de interacción que les permite relacionarse con el entorno natural y social de maneras que no están dirigidas exclusivamente a mantener vivo al organismo.

Froese y Di Paolo argumentan que este tipo de "normatividad relacionada a otro(s)" es un requisito para la cognición social, es decir, para ser un agente social (Froese y Di Paolo, 2011, p. 17). De este modo, los autores proponen la noción de una forma específica de agencia capaz de constituir un dominio cognitivo sensorio-motor: el *agente cognitivo*; y sólo un agente cognitivo puede ser un agente social.

Ahora bien, el resultado de esta relativa independencia respecto a los procesos metabólicos, es que la estabilidad de una estructura cognitiva autónoma depende tanto de la actividad electro-química del sistema nervioso, como también, de la forma en que tal estructura se acopla a los ciclos sensoriomotores, es decir, en el caso humano, las interacciones del agente con el entorno ya no son sólo para obtener todo aquello que le permite mantenerse vivo (ej. el alimento), sino que ahora es capaz de generar patrones de percepción-acción impulsados por valores sociales dirigidos a otros, lo que le permite adquirir repertorios sensorio-motores básicos que configuran paulatinamente el patrón de actividades que constituyen su vida y, en última instancia, su identidad socio-cultural.

En particular, esto desvincula a la normatividad sensorio-motora de los estrictos confines de la normatividad biológica, debido a que ya no depende de los procesos metabólicos para efectuarse, sino que ahora es impulsada por valores sociales dirigidos-a-otros. Así, sólo un agente cognitivo, que es capaz de regular sus ciclos sensoriomotores de esta manera no-metabólica, puede dar lugar a un dominio social que es identificado por su normatividad específica. A continuación, explico el paso del agente sensoriomotor al agente normativo social, para lo que es importante reiterar las condiciones requeridas que debe cumplir el agente para que surja la normatividad social:

-Normatividad social: en este nivel, el agente social debe ser capaz de 1) crear valores no-metabólicos y relacionados a otro, lo que da lugar a, 2) que sea capaz de percibir al otro como agente autónomo, con su propio punto de vista, sus propias intenciones, sus propios sentimientos, etc⁵⁵. Una vez identificadas estas dos condiciones, ahora corresponde resaltar que, para hablar de interacciones genuinamente sociales, es requisito cumplir con los siguientes dos criterios:

- Criterio 1. En el desarrollo de la interacción, la autonomía de los agentes no debe ser destruida (este criterio debe cumplirse para que una interacción se pueda llamar efectivamente "social").⁵⁶
- Criterio 2. En la interacción, el *otro* no debe ser percibido simplemente como una herramienta, objeto o problema para ser resuelto por la habilidad cognitiva de un solo agente, *i.e.* no debe ser percibido sólo como parte del entorno físico, sino como *otro agente* autónomo, precario, vulnerable, con su propia perspectiva del mundo, sus propias preocupaciones, sentimientos, etc.

Criterio (1) también es requerimiento para la noción de hacer-sentido participativo (PSM). Criterio (2) no es un requerimiento necesario para PSM, pero sí para la interacción socio-cognitiva, y refiere a la exclusión de las interacciones en las que el agente cognitivo no percibe al *otro* como agente autónomo y sólo lo encuentra como parte del entorno físico. Este criterio, desde la perspectiva clásica de la moral, significaría no considerar al otro de manera instrumental.

Recordemos que el enactivismo otorga crucialmente autonomía al proceso de la interacción, por lo que las interacciones dan lugar a estructuras autónomas que habilitan y restringen la conducta individual de los participantes. Sin embargo, no en todos los casos se les puede llamar "interacciones sociales", a menos que se cumplan criterios (1) y (2). Tales situaciones hacen uso de una normatividad social que es enfáticamente vinculada a *otros*, lo que constituye un dominio cognitivo *relacional* que es genuinamente social.

⁵⁵ Aquí se puede ver una similitud con el principio de *imparcialidad* kantiano, es decir, desde una perspectiva del lenguaje moral clásico, esta idea parecería acercarse a dicho principio en donde los intereses del otro importan tanto como los míos. La *imparcialidad* kantiana requiere que reconozcamos que todos los individuos son (moralmente) importantes y que no podemos descartar grupos de personas simplemente porque no tienen una relación personal con nosotros (ver Bramer, 2010, p. 137).

⁵⁶ Se verá que este criterio habilita la dimensión de la normatividad moral enactiva.

Así, desde este punto de vista, las interacciones que cumplen los criterios (1) y (2) y que por ende emplean una “normatividad relacionada a otros”, constituyen un dominio propiamente social. En esta línea, el agente sensorio-motor pasa a ser un agente normativo social si cumple los requerimientos mencionados y en el proceso de la interacción se satisfacen estos dos criterios.

Ahora bien, por último, la normatividad social así definida es, en naturaleza, *co-regulada* y *situada*. Es *co-regulada* debido a que la regulación implicada en la interacción socio-cognitiva es una regulación que crea una apertura para una acción que debe ser complementada o completada por otro agente, por lo que el éxito de la regulación depende esencialmente de una coordinación adecuada entre los agentes participantes (Froese y Di Paolo, 2011), es decir, exige *mutualidad* entre ellos. Un ejemplo que ilustra de manera sencilla este punto es el acto de "dar", el cual es completado exitosamente si el otro agente ejecuta la acción de "recibir" (2011, p. 24). En suma, esta normatividad apunta a las acciones *co-reguladas* y efectuadas conjuntamente entre los agentes participantes.

La razón es que durante el desarrollo de la interacción la regulación es coordinada interactivamente entre los agentes participantes, lo que da a la situación una “cualidad social” que exige mutualidad y que otorga el rol de *participantes* a los interactuantes (la “cualidad social” refiere a que una acción debe ser completada o complementada por otro agente). Entonces, la normatividad social enactiva es una regulación *co-dependiente* (o, en breve, es una *co-regulación*) que es sostenida momento a momento por los agentes que interactúan. Esta última característica hace de la normatividad social, una normatividad *situada*.

Es *situada* debido a que se centra en la constitución momentánea de las normas que emergen dinámicamente durante el proceso de la interacción, es decir, surge de ésta misma y ocurre momento a momento, habilitando y restringiendo los comportamientos de los agentes en un acoplamiento mutuo, en consecuencia, la normatividad que se desprende de la interacción socio-cognitiva no es algo predeterminado ni depende de un mecanismo fijo.

Así, pues, como vemos, se puede comenzar a hablar de la emergencia de la normatividad social (enactiva) mediante la interacción en segunda persona. Sin embargo, esto no es suficiente para capturar la idea completa de lo social y, por tanto, de lo moral. Para ello, se debe ampliar dicha definición de forma que permita ir más allá de sus estrictos confines,

para lo que es útil contemplar un orden normativo más abarcador, *i.e.* el orden normativo moral propio del entorno socio-cultural, tema que abordo en la siguiente sección.

En suma, hasta aquí he desglosado la definición operacional de la *cognición socio-cognitiva*, la cual posee tres características clave que recapitulo a manera de consecuencias de este apartado.

3.4.3. Consecuencias

1. La cognición socio-cognitiva es el acoplamiento sensoriomotor co-regulado que se da entre dos o más agentes cognitivos capaces de generar valores o normatividad (no-metabólica) relacionada a *otros*, lo que les permite ser capaces de percibir al *otro* como un agente autónomo en su propio derecho.
2. Para que una interacción sea efectivamente social, la autonomía de los agentes no debe ser destruida en el proceso.
3. De ello, emerge una normatividad *co-regulada* y *situada*, pues ésta surge momento a momento y exige que la regulación sea coordinada entre los agentes participantes en la interacción⁵⁷, lo que da lugar efectivamente a la normatividad social enactiva.

Pero, ¿qué la dota de matices morales? En un marco enactivo, la respuesta viene de dos puntos: 1) de la experiencia moral que se da en el encuentro con el otro agente y 2) de la normatividad moral propia del contexto. Estos puntos constituyen la normatividad moral enactiva (Pescador y Mojica, 2022), tema que desarrollo en el siguiente apartado.

3.5. Cognición moral enactiva

En esta sección, primero abordo la propuesta de la normatividad moral enactiva y, en segundo lugar, presento la evidencia empírica que apoya los aspectos corporizados involucrados en las facultades morales. Esto debido a que la explicación normativa permite aproximarnos a las condiciones corporizadas que habilitan la cognición moral.

3.5.1. Normatividad moral enactiva

El enactivismo se encuentra en una etapa incipiente respecto al estudio de la cognición moral humana, por lo que en co-autoría con Laura Mojica, proponemos una explicación enactiva acerca de la normatividad moral (Pescador y Mojica, 2022). En este apartado me

⁵⁷ Estas interacciones coordinadas, desde un enfoque moral clásico, en realidad podrían explicarse en términos de interacciones cooperativas (Sterelny, 2013).

baso en dicha propuesta, lo que permite develar algunos aspectos corporizados de la cognición moral.

Como punto de partida, nos basamos en la explicación *interaccionista* de Colombetti y Torrance (2009), arguyendo que la normatividad moral surge de la interacción misma, pero vamos más allá al proponer que se instituye en las prácticas socio-culturales.

Concretamente, la propuesta es que la normatividad moral enactiva emerge dinámicamente de la interacción social, debido a que ésta es la *condición de posibilidad* para cambiar irrevocablemente la autonomía de los agentes (positivamente o negativamente); y, en esta línea, identificamos dos condiciones para sostener que se instituye en la práctica social: 1) la naturaleza corporizada, precaria y vulnerable de los agentes, lo que resulta en una necesidad intrínseca de interactuar continuamente con otros y con su entorno (*dependencia constitutiva*) y; 2) su capacidad socio-lingüística que les permite expandir sus posibilidades y distinguirlas normativamente.

El argumento es como sigue: por un lado, proponemos un principio moral básico con valor universal y transcultural basado en los fundamentos enactivos, a saber, *el cuidado de las autonomías* de los agentes, mismo que no es susceptible de apropiación cultural. Por otro lado, aceptamos cierto relativismo moral respecto a las distintas formas culturales en las que se lleva a cabo tal principio, esto debido a que aceptamos un pluralismo de realización de dicho principio. La propuesta intenta hacer justicia a la naturaleza interactiva y situada de la moral, permitiendo contar con una herramienta evaluativa que sirva para discernir entre las acciones, prácticas y culturas opresivas de aquellas que promueven y cultivan la vida autónoma de los agentes.

En este punto, se debe agregar que dicho principio enfocado al *cuidado de las autonomías* puede ser consistente con los principios morales pre-lingüísticos computacionalistas, puesto que plausiblemente puede ser parte de aquellas orientaciones normativas generales, incontrovertidas y compartidas universalmente. En adelante, esto se verá más claro.

Ahora bien, para desarrollar la propuesta, comienzo por definir los tres niveles normativos que distinguimos para dar lugar a la normatividad moral enactiva: 1) el nivel de una *normatividad fenomenológica* que resulta en un *tirón moral* experimentado a nivel de primera-persona, lo que permite a los agentes conectarse afectiva y normativamente; 2) el

nivel de la *interacción entre agentes*, caracterizado principalmente como la *condición de posibilidad* para cambiar irrevocablemente la autonomía propia y ajena y, 3) el *orden normativo moral de fondo (background)* relativo al contexto y a la cultura en la que tienen lugar tales interacciones.

Crucialmente, el nivel (3) moldea heteronormativamente los otros dos niveles, y el nivel (2), de la interacción, moldea *heteronormativamente* el nivel fenomenológico de los participantes (*i.e.*, el nivel 1). Así, la propuesta resalta la importancia de las autonomías individuales, pero también el papel de las heteronomías sociales y su interrelación (2022, p.17). La normatividad moral así definida sigue la línea enactiva, por lo que, desde este punto de vista, tal normatividad es fundamentalmente corporizada, situada, afectiva y pre-reflexiva, y, de manera relevante, emerge de la interacción y surge como una práctica social.

En torno a ello, sugerimos que, debido a la condición natural de vulnerabilidad de los agentes humanos y de su dependencia social constitutiva, la normatividad moral enactiva surge de las interacciones, puesto que en éstas emerge *la posibilidad* de cambiar irrevocablemente la autonomía del otro, ya sea de manera positiva o de manera negativa y, por tanto, los agentes son moralmente responsables de cuidar la autonomía propia y ajena en la interacción (atributo habilitado por el carácter de *participante*⁵⁸ que otorgan las explicaciones de la cognición social enactiva, las cuales, también destacan las restricciones propias de la interacción).

Ahora bien, *cuidar la autonomía* de los agentes significa enfáticamente que en las interacciones “se deben garantizar las posibilidades para que los agentes afirmen y desarrollen sus propias identidades” (2022, p.2), tanto físicas como socioculturales, lo que se postula como la exigencia normativa mínima que regula la conducta individual y las interacciones sociales. Siguiendo una coherencia conceptual enactiva, esta normatividad es *situada* en interacciones concretas y encuentra una justificación pragmática y afectiva.

⁵⁸ Al respecto, van Grunsven (2018) ha sostenido que las interacciones corporizadas en segunda persona o, más bien, la *experiencia* de los involucramientos en segunda persona da lugar a nuestra orientación de ser *participantes* en las interacciones que sostenemos, lo que, según la autora, implica tener responsabilidad por el otro. En otras palabras, la perspectiva *experiencial* de las interacciones en segunda persona implica una orientación de participación, a partir de lo cual, Grunsven ha sugerido que dicha noción de participación implica responsabilidad hacia el otro. Así, desde este punto de vista, los participantes de cada interacción, tendrían responsabilidad por el otro (ver van Grunsven, 2018).

De manera relevante, nuestra propuesta se basa en el requerimiento básico del PSM y de la interacción socio-cognitiva, a saber, que “la autonomía de los agentes no debe ser destruida en la interacción”, lo que implica la *posibilidad* de cambiar irrevocablemente la autonomía de los interactuantes. En nuestra interpretación, esto da lugar a la dimensión moral, pues las autonomías precarias y vulnerables de los agentes se encuentran permanentemente expuestas en la interacción, así que a cada momento de la misma existe la *posibilidad* de aumentar, cambiar o dañar la autonomía propia y ajena. De ahí que los agentes sean moralmente responsables de cuidar y respetar la autonomía de los participantes en toda interacción social⁵⁹.

A continuación, explico cada nivel normativo que constituye la normatividad moral enactiva que proponemos.

Como punto de partida, tomamos las condiciones que habilitan la interacción socio-cognitiva tal como se definió: éstas son interacciones en 2ª persona que exigen mutualidad, en donde los agentes que interactúan desempeñan el rol de *participantes* (y no de meros observadores) y ellos deben ser capaces de crear valores relacionados a otros agentes, característica que les permite percibir al *otro* como agente autónomo en su propio derecho. Esto da lugar a la normatividad social, la cual es fundamentalmente co-regulada y situada, tal como se vio.

La interacción socio-cognitiva así definida implica que en cada momento de la interacción la autonomía de los agentes participantes esté expuesta, característica que implica grados de moralidad.

En torno a ello, proponemos que la normatividad moral emerge también de otros dos niveles: del nivel fenomenológico, experimentado en 1era persona, que se manifiesta como un *tirón moral* y es exclusivo de la experiencia (el *tirón moral* representa un tirón fenomenológico que *experimentamos* cuando se nos presenta la alteridad, pues la experiencia subjetiva específica de “otra persona” no ocurre con los objetos inanimados), y, por otra parte, del orden normativo moral relativo al contexto y a la cultura, que es más abarcador que los otros dos niveles. Así, cada nivel (el individual, el de la interacción y el del contexto socio-cultural) posee su propia autonomía, pero a su vez, son heterónomos

⁵⁹ El deber moral recae en el hecho de la permanente exposición de nuestra autonomía vulnerable en las interacciones.

con respecto a los otros niveles: el nivel individual es habilitado y restringido por la interacción y la interacción a su vez es regida por el contexto socio-cultural. Veamos esto con más detalle.

Comienzo con el nivel de la normatividad fenomenológica. Se debe tener en cuenta que este nivel normativo refiere únicamente a un nivel experiencial.

3.5.2. Normatividad fenomenológica

Siguiendo a Sophie Loidolt (2018), nos enfocamos en un nivel específico de normatividad que surge de la estructura experiencial que se da en *el encuentro con el otro* y tiene dos características constitutivas: 1) surge de los involucramientos con otro agente, y, 2) confronta al agente con un "deber". Nos enfocamos en este tipo de normatividad para caracterizar la experiencia moral que contribuye a la conexión afectiva y normativa entre los interactuantes.

Primero, en la tradición fenomenológica (ej. Scheler, Husserl, Heidegger) se ha argumentado que *el encuentro con el otro* nos confronta con un "deber", y este "deber" ha sido concebido como un *llamado* a la conciencia, como un *llamado* a responder con *vocación, i.e.*, a responder con amor (en la ética de Husserl y de Scheler, respectivamente), o bien, como un *llamado* a la responsabilidad (en Levinas). El planteamiento es que, a nivel experiencial, somos receptivos a ese *llamado* del otro, somos receptivos a los mandatos de ese "deber", debido a que emerge una estructura experiencial que involucra tres elementos: los valores, el yo y el otro. Esta estructura experiencial instituye este nivel normativo (Loidolt, 2018).

Ahora bien, este nivel de normatividad es constitutivo de la experiencia moral subjetiva y se hace presente en el "encuentro con el otro", el cual se siente en el cuerpo vivo como un *tirón moral*, que nos permite abrirnos afectiva y normativamente con los demás. Este *tirón moral* representa un tirón fenomenológico que experimentamos cuando se nos presenta la alteridad, pues la experiencia subjetiva específica de "otra persona" no ocurre con los objetos inanimados. De esto, deriva un nivel normativo fenomenológico específico del encuentro con otro agente, lo que no pretende ser un ideal ontológico o metaético, sino, únicamente resaltar lo que ocurre a nivel experiencial en el encuentro entre agentes; no percibimos al otro como un cuerpo físico desprovisto de significado. Esto contribuye a conectarnos afectiva y normativamente, lo cual es crucial para interactuar moralmente, al menos en nuestra propuesta (2022, p. 11, nota al pie 11).

Crucialmente, en la fenomenología la normatividad se entiende como proveniente del sujeto experimentador, pero al mismo tiempo, del objeto experimentado, dando lugar así a la idea del aspecto *relacional* de la normatividad moral y por tanto de los valores morales. Esto es relevante debido a que, en las teorías éticas tradicionales, tales como la humeana o la kantiana, se ha establecido que los valores morales se encuentran, o en el sujeto, o “afuera” en el mundo, ignorando la posibilidad de un aspecto *relacional* de los mismos (ver van Grunsven, 2015). En contraste, en una explicación enactiva se debe destacar el aspecto *relacional* de la normatividad, debido a que *la presencia del otro* (real o virtual) es condición para una normatividad moral enactiva.

En este sentido, la experiencia perceptiva del otro como persona autónoma, precaria y vulnerable, y de nosotros mismos como tal, se forma, enriquece e incluso distorsiona en la propia interacción (van Grunsven, 2015), de tal manera que, a nivel experiencial, somos *llamados* mínimamente a respetar y a cuidar la autonomía propia y ajena en cada momento de la interacción, debido a que nuestra condición natural de vulnerabilidad nos mantiene en continua exposición. Por consiguiente, desde este punto de vista, percibir al otro como tal en una interacción y la normatividad que de ella emerge, no son elementos pre-dados ni tampoco dependen de un mecanismo fijo, por el contrario, constituyen una condición que se negocia continuamente en el proceso de la interacción social⁶⁰.

Así, los interactuantes tienen la *posibilidad* de explorarse y retroalimentarse mutuamente a cada momento de la interacción. Si tal mutualidad es positiva o negativa depende de la responsabilidad (moral) de cada uno de ellos. De ahí que se abra un amplio abanico de posibilidades en las que los valores morales entran en juego desde el inicio de toda interacción social (esto debido a la vulnerabilidad intrínseca de los agentes y de la dependencia recíproca que hay entre ellos). Sin embargo, ¿cómo emergen los valores morales? ¿de dónde provienen? ¿llegan únicamente de ese *tirón moral*?

Para resolver estas preguntas, en la siguiente sección, el foco es un nivel normativo más abarcador, a saber, el orden normativo moral propio del contexto, o bien, la normatividad moral socio-cultural que tiene autonomía en su propio derecho, con lo que se hace justicia al nivel supra-personal desde una metodología relacional.

⁶⁰ Es preciso recordar que aquí estoy revisando la teoría enactiva de manera *purista*, con la finalidad de revelar en qué aspectos se enfoca y qué explica, es decir, no es que se contradiga lo visto en el computacionalismo del capítulo anterior.

3.5.3. Normatividad moral socio-cultural

Comienzo por establecer que un distintivo de la cognición humana y de las interacciones humanas en general, es que siempre se despliegan en un contexto cultural que implica un orden normativo de fondo (Froese y Di Paolo, 2011). Se trata de un orden pre-existente de prácticas socio-culturales compartidas que involucran valores históricos heredados y una parte constitutiva de esa gran parcela normativa son los valores morales. Al respecto, recordemos que, en esta línea, la distinción de los valores morales es que ellos están asociados con las formas en que se efectúa el *cuidado de las autonomías* de los agentes.

Ahora bien, decir que hay valores culturales (o morales) que guían nuestra conducta y entendimiento apunta a un fenómeno más general, ya que el proceso de enculturación tiene efectos profundos en la constitución de nuestro comportamiento individual. Lo mismo ocurre con la moralización humana, debido a que hay un impacto constitutivo de los valores morales en cada etapa de nuestra formación y por consiguiente en nuestro comportamiento. Por lo tanto, la moralización implica, al menos, alguna forma de *heteronomía* normativa que constituye parcialmente el comportamiento de los agentes y de las interacciones.

Pero ¿cómo se relacionan *heteronormativamente* las interacciones socio-cognitivas concretas con esa normatividad moral socio-cultural?

El punto de partida, es tomar en cuenta que una implicación crucial de la propuesta interactiva de Colombetti y Torrance (2009) es el cambio de la unidad de análisis: del *individuo* al dominio de la *interacción* y, con ello, se cambia del aspecto *individual* al aspecto *relacional*; movimiento que impacta profundamente en el estudio de la moral, puesto que, lo que cada uno de nosotros hace en relación con *otro* es parcialmente estructurado y caracterizado en términos interindividuales e interpersonales. Se puede decir que, desde esta perspectiva, el carácter moral de una situación dada surge (al menos en parte) de los significados compartidos que emergen de las interacciones entre los participantes.

Por consiguiente, aquí me enfoco en la dimensión *relacional* del nivel supra-personal, es decir, de las interacciones socio-cognitivas que se dan entre dos o más agentes cognitivos y de la estructura normativa moral que forma parte del contexto socio-cultural.

La estructura normativa moral socio-cultural es un distintivo único de las interacciones humanas (siendo éste un subconjunto de las mismas, pues claramente hay muchos otros tipos de normas). La teoría sociológica señala que las acciones de los agentes humanos

tienen lugar en un contexto de prácticas compartidas, de normas sociales, reglas, expectativas, etc. que circunscriben fundamentalmente lo que los agentes hacen en cada etapa de su formación, es decir, constituyen, constriñen y habilitan la forma en que las personas actúan; característica que ofrece una gama de posibilidades para la acción social en correspondencia con el contexto (ver Steiner y Stewart, 2009). Estas normas sociales más amplias incluyen un carácter fuertemente moral por estar asociadas a las maneras en que el *cuidado de las autonomías* debe llevarse a cabo y, por tanto, contribuyen a determinar el carácter moral de las interacciones socio-cognitivas que comprenden la actividad concreta de nuestra vida social cotidiana.

En este aspecto, la definición de la interacción socio-cognitiva⁶¹ en conjunto con la noción de *heteronomía* propuesta por Steiner y Stewart (2009) -la cual veremos a continuación-, permite ampliar el término de autonomía. Se trata, pues, de explicar los impactos directos que tiene una normatividad heterónoma en la interacción social y en la conducta individual, lo que involucra elementos constitutivos, autónomos y heterónomos.

- **Heteronomía**

En breve, Steiner y Stewart (2009) proponen la noción de *heteronomía* como un estado o actitud propia de una persona o un grupo para el acatamiento de un orden normativo socio-cultural pre-establecido constituyente del contexto, el cual dicta los estándares compartidos que determinan si los comportamientos o las interacciones son moralmente correctas o no. Este nivel normativo se compone de criterios externos que van más allá de las impresiones individuales y de las interacciones concretas, lo que permite determinar si los comportamientos, prácticas o interacciones en ese contexto son correctas o incorrectas, elogiables o reprobables, opresivas o liberadoras, etc.

En este sentido, se trata de hablar de una normatividad más abarcadora que contribuye a conformar tanto la experiencia en primera persona como el nivel de la interacción, puesto que este ámbito normativo impone restricciones importantes sobre cómo se desarrollan nuestras acciones e interacciones en la vida cotidiana, de tal manera que, la existencia y la ubicuidad de tales restricciones permiten que hablemos de este campo normativo como uno de *heteronomía* más que de *autonomía*.

⁶¹ Considérese que la interacción socio-cognitiva no toma en cuenta el papel fundamental que desempeñan las normas sociales y morales del contexto en el que tienen lugar, por ello se busca explicarlo desde la noción de heteronomía de Steiner y Stewart (2009).

El concepto de *heteronomía* refiere, entonces, a que un agente genere un comportamiento que respete y acate las normas del contexto, en donde dichas normas son habilitadoras y restrictivas, pues nos permiten encontrarnos y coordinarnos, así como también constituyen la posibilidad de producir nuevos comportamientos. Así, ser *heterónimo* es una actitud que los agentes adoptan implícitamente en su comportamiento diario, lo que se ejemplifica constantemente en sus acciones. Debido a que un determinado comportamiento se produce de acuerdo con las normas compartidas, constituye una acción y, por tanto, adquiere sentido para otros agentes (Steiner y Stewart, 2009).

Entonces, de ello, se puede decir que las interacciones socio-cognitivas son regidas, en parte, por ese orden normativo socio-cultural y moral propio del contexto en donde tienen lugar, pues constituye una fuerza normativa moldeadora que instaaura, guía y evalúa los comportamientos y las interacciones a partir de estándares compartidos.

Ahora bien, desde la noción de heteronomía, Steiner y Stewart proponen que un ámbito genuinamente social se constituye en gran medida por este orden normativo preexistente y culturalmente heredado en el que cada agente social se encuentra incrustado a lo largo de su vida diaria. Las normas en esta estructura preexistente "en realidad constituyen la posibilidad de *enactuar* o de *hacer-sentido* de mundos que simplemente no existirían sin ellas. [Por lo que] las interacciones entre dos o más agentes nunca son propiamente sociales, a menos que tengan lugar en el contexto de un entorno de estructuras o normas sociales que den sentido a las interacciones" (Steiner y Stewart 2009, p. 528). Es decir, este orden normativo proporciona en gran medida los medios para que se produzcan las acciones e interacciones y que éstas adquieran sentido para los demás.

En comparación, recordemos que en la interacción socio-cognitiva el ámbito social emerge de la dinámica autónoma de la interacción a medida que se desarrolla en el aquí y el ahora, por lo que su concepción normativa es más limitada y corresponde a las necesidades concretas de la situación. No obstante, se verá que tales interacciones en realidad son "heterónomas" con respecto a esas estructuras preexistentes a gran escala.

Pero, específicamente ¿cómo se relacionan las interacciones socio-cognitivas concretas con la normatividad moral del contexto socio-cultural? La respuesta radica en la genuina complementación de la *autonomía* y de la *heteronomía*; revisemos cómo podría efectuarse este movimiento.

- **Autonomía y heteronomía como procesos complementarios**

La noción de interacción socio-cognitiva refiere a una situación dada entre dos o más agentes autónomos, como una dinámica de interacción independiente que tiene su propia autonomía y que, a su vez, constriñe las actividades de los participantes individuales en la situación. Sin embargo, claramente, en la medida en que los participantes individuales en la interacción están "sujetos" a esta dinámica, son "heterónomos" con respecto a la dinámica misma. Por lo tanto, la heteronomía podría verse como una característica propia de la interacción socio-cognitiva (Torrance y Froese, 2011), puesto que el comportamiento individual es heterónimo con respecto a la interacción.

Ahora bien, en relación con las estructuras normativas morales preexistentes, la dinámica de la interacción puede ser descrita como heterónoma en lugar de autónoma. Pero al mismo tiempo, en relación con los participantes individuales, este proceso interactivo ocupa un grado de autonomía, como se describe en el enfoque de la interacción socio-cognitiva.

Asimismo, el orden normativo moral preexistente, al que hace referencia la noción de heteronomía, puede describirse igualmente como "autónomo" (independiente, trascendente) con respecto a los participantes, pues configura un sistema autónomo en sí mismo en la medida en que produce los procesos mismos que le constituyen, a saber, el patrón de interacciones compartidas por la comunidad.

En este contexto, "autonomía" y "heteronomía" no se contraponen y en cambio se pueden entender como términos relacionados. En ambos casos hay agentes individuales y una estructura supraindividual. En cada caso, la estructura supraindividual (interacción dinámica en el enfoque de interacción socio-cognitiva y orden normativo históricamente dado en el enfoque de heteronomía) tiene "autonomía" en el sentido amplio de ser "independiente", de tener "vida propia". Pero también, en ambos enfoques, esta estructura restringe y habilita las acciones de los individuos, y de esta manera los agentes individuales son heterónomos con respecto a la estructura general, debido al papel constitutivo y habilitador que la estructura normativa moral tiene en su actividad.

Así, las interacciones socio-cognitivas adquieren cierto carácter moral debido al papel constitutivo que desempeñan las normas morales pre-establecidas⁶², pero también, esas

⁶² Por supuesto, muchas interacciones son neutrales, aquí solo hablamos de la *posibilidad* que tienen de ser morales.

normas morales pre-establecidas se perpetúan a través de interacciones continuas entre los miembros de la comunidad para quienes esas normas morales tienen fuerza.

En este sentido, las acciones e interacciones humanas toman lugar dentro de un contexto histórico permeado por la normatividad moral enfocada en el *cuidado de las autonomías* de los agentes, de tal manera que en el momento de la acción o la interacción en sí, existe la dinámica de cómo los agentes participantes son moldeados parcialmente por estas condiciones normativas en su interacción, pero también, las acciones de cada agente individual son moldeadas por las acciones de los otros agentes participantes en esa situación. Por tanto, una aproximación enactiva al estudio de la moralidad necesita combinar este dominio de la dinámica interindividual con las condiciones normativas morales del contexto.

Ahora bien, tras revisar estos argumentos, a continuación presento puntualmente la propuesta de la normatividad moral enactiva.

3.5.4. Normatividad moral enactiva: la propuesta condensada

Hasta aquí he presentado: 1) la definición operacional de la interacción socio-cognitiva con sus componentes, 2) la normatividad fenomenológica a partir de la experiencia moral que emerge del encuentro con el otro (*tirón moral*), y, 3) expliqué la relación entre las interacciones socio-cognitivas y el orden normativo moral socio-cultural, lo que se explica gracias a que la autonomía se amplía por la heteronomía, en donde hay elementos heterónomos y de constitución (para ilustrarlo, ver Fig. 1).

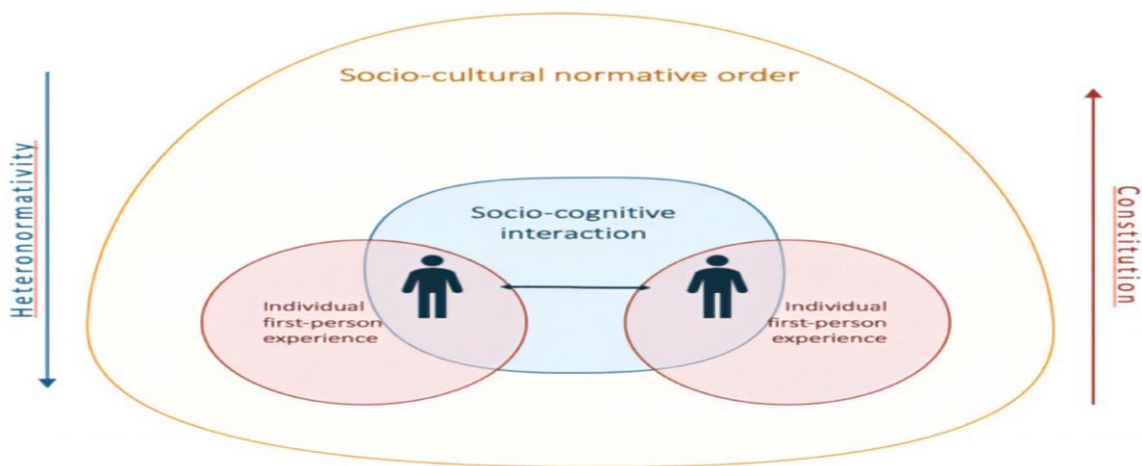


Fig. 1. La flecha izquierda captura dos relaciones heterónomas de conformación y condicionamiento normativo: el orden normativo sociocultural que constituye el entorno de los individuos da forma y regula heterónomamente las interacciones humanas, así como las interacciones sociocognitivas dan forma y regulan heterónomamente las experiencias y comportamientos del individuo. Por otra parte, la flecha derecha captura dos relaciones de constitución: el patrón de interacciones y el patrón de comportamiento de los individuos constituyen el orden normativo sociocultural que sirve como contexto socialmente compartido. Estos patrones se extienden en el tiempo y a través de múltiples individuos, y constituyen la forma de vida compartida (las interacciones sociocognitivas articuladas están constituidas por los comportamientos individuales de dos o más agentes que se coordinan en el tiempo y el espacio de tal manera que se reconocen mutuamente como agentes con sus propios fines e intereses y no destruyen las autonomías de cada uno). (Esquema extraído de Pescador y Mojica, 2022, p.14).

De ello, el argumento puede instanciarse de la siguiente manera:

Premisa: en la interacción (socio-cognitiva) se abre la *posibilidad* de dañar o aumentar la autonomía propia y ajena, lo que instaura una dimensión moral *relacional* en la que los interactuantes adquieren la responsabilidad moral (mínima) de cuidar y respetar mutuamente sus autonomías (es moral por la *posibilidad* que existe de cambiar la autonomía propia y ajena que surge de la interacción); proceso en el que distinguimos los siguientes tres niveles normativos con sus específicas características:

1) El nivel de la interacción (socio-cognitiva): se da entre dos o más agentes capaces de generar valores relacionados a otros, lo que les permite percibir al otro como tal, como un agente autónomo. La normatividad social que emerge y rige tales interacciones (en marcha) es una normatividad situada y co-regulada.

2) El nivel de la normatividad fenomenológica: en *el encuentro con el otro*, emerge un *tirón moral* fenomenológico que corresponde a la normatividad que surge de la experiencia específica de tres categorías interrelacionadas (los valores, el yo y el otro), lo que contribuye a la conexión moral y profundamente afectiva que se produce en el encuentro con el *otro*.

3) El nivel de la normatividad socio-cultural: tanto la experiencia en primera persona como las interacciones socio-cognitivas son habilitadas y restringidas *heteronormativamente* por el orden normativo moral de fondo (enfocado en *el cuidado de las autonomías*). Pero también, dicho orden normativo se constituye a través de las continuas interacciones extendidas en el tiempo y de las prácticas compartidas, lo que le otorga *autonomía* en toda regla.

Así, pues, en suma, las principales características que habilitan nuestra propuesta son las siguientes:

En la interacción socio-cognitiva se abren grados de moralidad, debido a que implica que los agentes, en su papel de *participantes*, deben responder a la mutualidad que la interacción exige, no sólo para construir significados compartidos, sino que genuinamente adquieren grados de responsabilidad para con el otro, en donde la exigencia mínima es el *cuidado de las autonomías* (propia y ajena).

La razón es que, en el proceso de la interacción, la naturaleza autónoma, precaria y vulnerable de los agentes humanos queda expuesta momento a momento, es decir, nuestra constitución corporizada como seres vivos precarios y por lo tanto vulnerables y necesitados de la interacción con los demás y con el entorno (dependencia recíproca), resulta en una condición de vulnerabilidad constitutiva, la cual implica que la *posibilidad* de cambiar irrevocablemente la autonomía propia y ajena siempre esté presente en las interacciones entre agentes humanos (aunque, por supuesto, hay interacciones que nos dejan intactos). Esto caracteriza la profunda responsabilidad moral que tienen los agentes en toda interacción.

De esta manera, identificamos las dos condiciones necesarias para que la normatividad moral surja como una práctica social, es decir, como un ámbito normativo que va más allá de las impresiones individuales y de las interacciones particulares: 1) nuestra constitución corporizada, precaria y vulnerable que resulta en esa dependencia recíproca, y; 2) nuestra naturaleza sociolingüística, que nos permite expandir exponencialmente nuestras posibilidades de acción e interacción y distinguirlas normativamente, en particular, en lo que se refiere a las posibilidades de mantener y expandir la autonomía propia y ajena en la interacción.

Ahora bien, proponemos *el cuidado de las autonomías* de los agentes como un principio moral básico con valor universal, el cual a su vez posee pluralismo de realizabilidad, puesto que las formas en que puede efectuarse tal principio son dictadas por cada cultura o comunidad, es decir, en qué consiste y cómo se lleva a cabo *el cuidado de las autonomías* es determinado en gran medida por las prácticas socio-culturales de la comunidad en cuestión, vinculadas con la situación concreta y la historia que subyace a una interacción.

Esto significa que el curso de acción moralmente correcto es contribuir a mantener y, en el mejor de los casos, a maximizar la autonomía propia y ajena⁶³, afirmación que constituye

⁶³ Recordemos que *cuidar la autonomía* de los agentes significa enfáticamente que en las interacciones “se deben garantizar las posibilidades para que los agentes afirmen y desarrollen sus propias identidades”

un criterio transcultural objetivo, pero a su vez, su contenido y significado es provisto por las prácticas socioculturales realizadas por la comunidad específica a la que se pertenece (Pescador y Mojica, 2022, p. 15).

Crucialmente, nuestra propuesta implica una distinción entre *carácter moral* y *contenido moral*. En breve, el *carácter moral* está implicado en las conductas, interacciones o prácticas que presentan la posibilidad de cambiar irrevocablemente la autonomía de los agentes, ya sea de manera positiva o negativa, y, el *contenido moral* refiere enfáticamente a la especificidad de lo que significa *el cuidado de las autonomías* para esa comunidad en cuestión.

De ello, podemos distinguir las acciones, prácticas y culturas que mantienen o amplían las autonomías de las que oprimen o limitan las potencialidades de los agentes, a la vez que podemos extraer su carácter meramente *situado*. Esto permite efectuar un análisis de grado que aborde tanto las circunstancias específicas y situacionales como los criterios generales y objetivos de los mismos.

Así, pues, la caracterización de la normatividad moral que proponemos se instaura a partir de los tres niveles mencionados (en donde hay elementos autónomos, heterónomos y de constitución, ver Fig. 1), lo cuales enlisto a continuación a manera de consecuencias.

3.5.5. Consecuencias

1. En el encuentro con el *otro*, los agentes experimentan un nivel moral normativo manifestado como un *tirón moral*, y esto implica un grado de moralidad que se siente en el cuerpo vivo.
2. En la interacción sociocognitiva existe una dependencia recíproca entre los agentes, lo que da lugar a la dimensión moral que corresponde a la responsabilidad de cuidar la autonomía propia y ajena.
3. Las prácticas socio-culturales de la comunidad en la que se desarrolla la interacción moldean heteronormativamente el *contenido moral* adecuado para el cuidado de las autonomías (Ibid., p.16).

(Pescador y Mojica, 2022, p.2), tanto físicas como socioculturales, lo cual se postula como la exigencia normativa mínima, que regula la conducta individual y las interacciones sociales. Por lo tanto, maximizar significa aumentar tales posibilidades

Ahora bien, tras ver la propuesta enactiva de la normatividad moral, ahora corresponde abordar el aspecto cognitivo corporizado de las facultades socio-morales y su evidencia empírica.

3.6. Facultades socio-morales: intervenciones corporizadas

Comienzo con una pregunta clave que permitirá ver las aportaciones específicas del enactivismo a la cognición socio-moral:

- ¿Qué resuelve empíricamente el enactivismo que el computacionalismo no?

En resumen, se verá que las contribuciones concretas son: 1) explica cómo hacemos sentido del otro como *otro*, lo que no es trivial, sino que es una característica constituyente de la moral; 2) explica cómo nos auto-individuamos como agentes autónomos; 3) a partir de (1) y (2), se obtiene una perspectiva des-centrada del mundo que forma parte de nuestra perspectiva moral, y; 4) explica la intencionalidad compartida y sus mecanismos subyacentes, lo que contribuye a formar una plataforma constituyente del desarrollo moral. Estas características derivan naturalmente de la metodología relacional que sustenta el enactivismo y cada una de ellas posee implicaciones cruciales para el desarrollo socio-moral.

Para ver esto de manera clara, la sección la divido en dos partes: primero, me enfoco en la evidencia de la “cognición social” y, posteriormente, me centro en la evidencia de la “cognición moral”, lo que me permitirá trazar puntualmente la intervención de las capacidades corporizadas en la cognición socio-moral, según mi propuesta.

Concerniente a la cognición social, son tres los conceptos básicos que reviso junto con sus implicaciones: 1) la perspectiva epistémica de 2ª persona; 2) la interacción social, y; 3) el involucramiento emocional.

La perspectiva de 2ª persona destaca la importancia del involucramiento emocional y de la interacción social debido a que estas características permiten *experimentar e interactuar* directamente con otros, lo que conforma nuestra forma básica de conocer a los demás. Es decir, esta perspectiva sostiene que (al menos en parte) conocemos *directamente* las mentes de los otros mediante la interacción y el involucramiento emocional, sin ningún otro mediador cognitivo como son las rutinas inferenciales o simulacionistas postuladas por la TT y por la TS (rutinas que algunos autores han llamado “desvíos intelectuales” por su costo

cognitivo). De esta manera, aquí se concibe que las rutinas atribucionistas son precedidas por estas capacidades corporizadas más básicas.

Ahora bien, concerniente a la cognición moral, me centro en la evidencia que muestra la capacidad de “intencionalidad compartida (mínima)” en infantes de 2 a 8 meses de edad. Esto permite ver las formas básicas de ayuda, de cooperación y de acciones conjuntas efectuadas sin representar la perspectiva del otro, sino más bien, catalizadas por la sensibilidad a la pertenencia de grupo y habilitadas por capacidades corporizadas; lo que sirve como andamio y facilitador de formas más sofisticadas y complejas de la moral.

Asimismo, expongo los tres mecanismos psicológicos que se han considerado la base de la intencionalidad compartida (mínima), los cuales son esencialmente corporizados: 1) conformismo social, 2) coordinación interpersonal básica, y, 3) cognición social no sofisticada. En conjunto, se plantea que estos elementos conducen a formas más complejas de acciones compartidas y de aprendizaje socio-moral en general.

Antes de examinar estos planteamientos, considero preciso presentar algunas implicaciones computacionalistas que encuentro problemáticas respecto a la cognición social.

3.6.1. Implicaciones problemáticas de la cognición social computacionalista

Como punto de partida, presento las cinco implicaciones que encuentro problemáticas en las explicaciones de la cognición social computacionalista y planteo, en contraste, cómo las aborda el enactivismo.

- Implicación 1: El *significado* de la acción humana se encuentra en los estados mentales, los cuales son inobservables.

En contraste, en el enactivismo experimentamos inmediatamente al otro como *otro* en la interacción, es decir, conocemos la mente de los otros mediante el involucramiento emocional en la interacción. Esto implica que somos *directamente* afectados por las acciones del otro y respondemos a ello en consecuencia; por tanto, las acciones de los demás conllevan *significado* en sí mismas, debido a que “la conciencia del otro es revelada en el acto mismo” (Asch, 1952, p.158).

La razón es que en el enactivismo el significado del comportamiento no está localizado en estados o procesamientos que portan información dentro del sistema (como estados

mentales ocultos), sino que el significado *surge* dinámicamente en la interacción misma, por tanto, en principio, las mentes de los otros y sus estados se nos presentan de manera directa y no son enteramente inobservables. Como se ha visto, en la interacción, mediante el involucramiento emocional, experimentamos inmediatamente al otro como *otro*, conocemos su mente y a su vez somos afectados y respondemos en consecuencia, por tanto, en las acciones de los otros percibimos sus significados, pues la conciencia se revela en la acción (Asch, 1952).

- Implicación 2: La percepción es *neutra* (la percepción representa características evaluativamente neutras del mundo), en tanto que la TT y la TS parten de la observación perceptual de datos físicos desnudos, *i.e.* de un comportamiento físico desnudo que “en principio carece de significado para el observador” (Gallese, 2009, pp. 520-21). La cognición social, desde este punto de vista, es la capacidad de representar este comportamiento desnudo y convertirlo “en algo que el observador sea capaz de comprender” (Gallese, 2009, p. 521).

En particular, esto contrasta con la “hipótesis de la percepción directa” (Froese y Leavens, 2014), pues las versiones computacionalistas suponen que los estados subjetivos de los demás son inaccesibles a la percepción y, por lo tanto, deben inferirse o simularse en función de la percepción de comportamientos superficiales puramente objetivos. Por ello, se considera que estos enfoques estándar no son fenomenológicamente convincentes (Froese et al., 2013). En contraste, desde el enactivismo, la percepción no es neutra, y el planteamiento es que conocemos al *otro* mediante el involucramiento emocional y la interacción misma, en donde no hay realmente un abismo epistémico entre el yo y el otro. Esto resulta en una visión del conocimiento social en donde el conocimiento de las otras mentes surge de la interacción *entre* agentes y no solo en la cabeza del agente individual.

En el enactivismo, un factor constitutivo de la percepción empática es que el agente experimenta el comportamiento corporal de los demás como expresión de su alteridad y de su propia autonomía: “el otro cuerpo no es un simple cuerpo, fragmento del mundo, sino que es encontrado como *otro* con su propia visión o perspectiva del mundo” (Grunsvén, 2015, p.225).

Esta comprensión empática del *otro* es fundamentalmente *asimétrica* (*i.e.* percibo al otro como *otro* agente autónomo y no como una proyección de mí mismo) y emerge de las interacciones recíprocamente acopladas entre agentes corporizados.

- Implicación 3: Las posturas atribucionistas no explican cómo hacemos sentido del otro como *otro* en su propio derecho. La razón es que tanto la TT como la TS parten de una perspectiva de mero observador (o espectador), por lo que a la TT se le acusa de ser afectivamente desvinculada (en particular, por su generalidad y por primar el uso de conceptos sofisticados); y a la TS se le acusa de ser meramente asimilacionista (en tanto que el entendimiento de los otros es una proyección de lo que yo he experimentado primero), entonces, esto deriva en que ninguna explica cómo hacemos sentido del otro como *otro* en su propio derecho.

En contraste, el enactivismo sostiene que hacemos sentido del otro como *otro*, es decir, percibimos genuinamente al otro como *otro* agente autónomo, no mediante la observación desvinculada, sino, mediante el involucramiento emocional en la interacción misma.

- Implicación 4: Los enfoques atribucionistas no explican adecuadamente la experiencia social, pues la TT y la TS parten de un punto de vista de observador y con ello distorsionan los hechos fenomenológicos de la cognición social, por lo que sus explicaciones carecen de descripciones detalladas de la experiencia social (Ratcliffe, 2007).

La experiencia social se da en la interacción misma, al experimentar una situación social mediante los involucramientos emocionales y la percepción apropiada de la estructura interaccional.

- Implicación 5: Las rutinas atribucionistas consideran que la interacción es la última etapa de la cognición social, es decir, según el modelo computacionalista de la cognición social, el agente adscriptor primero observa la conducta manifiesta del *otro*, después aplica la rutina atribucionista y por último interactúa.

Esto es problemático debido a que implica una distinción ontológica mente-conducta, de lo que surgen acciones desprovistas de significado y un abismo epistémico entre el yo y el otro. En contraste, el enactivismo plantea que la interacción misma constituye el fenómeno social, pues en el proceso de la interacción: i) comprendemos al otro como *otro*, ii) surgen los significados de sus estados y de sus acciones y, iii) respondemos a ello en consecuencia. Esto es, mediante la interacción conocemos y respondemos a las otras mentes, puesto que ejercitamos nuestro saber práctico de nuestro conocimiento social de los demás a través de las continuas interacciones. De hecho, hay una circularidad epistémica: entre más interactuamos con otras mentes, más las conocemos y más hay con

qué interactuar (Reddy, 2008), por lo que la interacción, desde este punto de vista, no puede ser la última etapa de la cognición social, sino, más bien, ésta constituye en sí misma la cognición social.

Así, la perspectiva de 2ª persona aborda la cuestión acerca de cómo reaccionamos afectivamente a los estados de los demás, mientras que la perspectiva de mero observador, *i.e.* de 3ª persona, parece enfocarse sólo en los procesos de explicación y predicción mediante la atribución mental, mas no especifica la cuestión acerca de cómo respondemos al otro. Si bien, en el proceso de atribución mental, el resultado (*output*) generado es para responder apropiadamente a la situación (a partir de la representación del otro y de sus estados intencionales), también es cierto que tal respuesta no apela a un sentido afectivo, es decir, no se explica la respuesta afectiva o emocional del atribuidor, en toda regla. En palabras de Bohl:

“Aunque por ej. TS explica cómo podemos atribuir automáticamente emociones a otras personas, no tiene en cuenta las propias reacciones emocionales del atribuidor como fuente de información sobre la situación social. Nuestras propias respuestas afectivas a menudo difieren de las emociones que reconocemos en los demás; por ejemplo, cuando reconozco la ira en el otro, puedo tener miedo. ToM puede explicar cómo reconozco la ira en el otro, pero no da cuenta de cómo me asusto como resultado” (Bohl y van den Bos, 2012, p. 4).

Esto se hace más claro si consideramos que el punto de partida de las rutinas atribucionistas es que la percepción, la cognición y la emoción son módulos distintos, de tal suerte que la percepción solo podría proporcionar hechos físicos neutrales, por lo que la cognición tiene que descubrir qué está pasando para después evocar una respuesta emocional (Froese, 2015, p. 3).

Así, pues, se trata de explicar los huecos que dejan las teorías atribucionistas a partir de la evidencia empírica que apuntala los principios enactivos de la cognición socio-moral.

3.6.2. Argumento

¿Por qué es relevante percibir al otro como *otro*? Hay al menos dos razones: 1) porque es un requerimiento de la interacción socio-cognitiva, y, 2) porque se ha establecido como un requisito para poseer una perspectiva genuinamente moral. Veamos esto con más detalle.

Comienzo con una afirmación hecha por Krosgaard (2010), quien defiende que para hablar de una perspectiva genuinamente moral es requisito ver al otro como *otro* en su propio derecho:

“[U]na parte esencial para superar la cosmovisión teleológica del animal es reconocer que las cosas no existen en relación conmigo. El mundo no... consiste en mis depredadores, mi presa, mi descendencia, ... Tener ese hecho firmemente a la vista es esencial para lograr... la concepción del mundo que hoy reconocemos como prácticamente racional, es decir, como moral. ... que [otros seres vivos] no existen para nosotros, y con referencia simplemente a nuestros intereses, sino que tienen una existencia independiente e intereses propios; comprender estos hechos es esencial para formar una concepción teórica de un mundo que existe independientemente de nosotros, así como una concepción práctica del mundo con el que debemos relacionarnos” (Korsgaard, 2010, pp. 33-4).

El planteamiento señala que se precisa reconocer a los demás como seres autónomos con “existencia independiente e intereses propios” (y no como meros fragmentos del entorno que sólo existen en relación conmigo) para lograr una perspectiva moral del mundo, lo que implica que nuestra visión moral del mundo depende de identificarnos a nosotros mismos y a los demás como seres autónomos e independientes.

Esto conduce a establecer una liga indisoluble entre la cognición social y la cognición moral, en donde la primera se identifica como condición necesaria de la segunda⁶⁴. No obstante, esto exige precisar la concepción de cognición social de la que partimos para explicar la moralidad.

En general, el término de “cognición social” refiere a nuestra capacidad para entender e interactuar con otros, lo que implica identificar a los demás como personas con sus propias necesidades, preocupaciones, creencias, deseos, intenciones, dolores, placeres, etc.

Sin embargo, debido a que el esquema atribucionista ofrece una imagen de vida según la cual estamos ante todo en el negocio de observar, explicar y predecir el comportamiento de otras personas, en donde la TT se presenta como “general y afectivamente desvinculada” y la TS como meramente “asimilacionista”, su concepción deriva en no ver al otro como *otro* en su propio derecho, sino como una herramienta o problema que requiere ser resuelto desde la cognición individual.

En contraste, el enactivismo muestra tener el potencial para explicar la cognición social en los términos requeridos, *i.e.*, en la interacción social los agentes hacen-sentido de sí mismos y de los otros como *otros* en su propio derecho. Esto debido a que la experiencia perceptiva de los otros es ante todo empática, puesto que, “en la empatía experimentamos a otro ser humano directamente como persona, es decir, como un ser intencional cuyos

⁶⁴ Nótese que esta liga indisoluble es un punto de convergencia entre el computacionalismo y el enactivismo.

gestos y acciones corporales expresan sus experiencias o estados mentales” (Thompson 2007, p. 386).

Este tipo de empatía es crucialmente *des-centrada*, es decir, hay una *asimetría* entre el yo y el otro, lo que permite identificar al otro como *otro* en su propio derecho y no como una mera proyección de mí. Esta característica *des-centrada* de la empatía es constitutiva de nuestra perspectiva humana sobre el mundo⁶⁵.

Pero, específicamente ¿qué es lo que permite individuar al otro como *otro*?

Para ser clara, respecto a las teorías atribucionistas, la idea de relacionarnos mediante TT o TS ya implica que se ha identificado al otro como a alguien a quién aplicar la atribución mental, pues le suponemos mente y racionalidad, pero esto no justifica verlo genuinamente como *otro* en los términos requeridos. La razón es que la naturaleza de la rutina atribucionista deriva en ver la mente del otro como un problema a resolver y justo por ello se le aplica la atribución, así que pensar que este requerimiento lo cumplen la TT y/o la TS ¿no socavaría el problema de las otras mentes? (Grunsvén, 2015).

En particular, el “problema de las otras mentes” proviene de su característica distinción ontológica mente-conducta, de lo que resulta la observación de una conducta o acción *sin* significado, de tal manera que la pregunta acerca de cómo realmente hacemos sentido de los otros se enmarca en esta metodología dualista (ver Costall et al. 2006; Leudar y Costall 2008):

“Cada uno es un observador independiente, en lugar de participar activamente con el otro en algún proyecto conjunto.

La información disponible para cada uno de nosotros sobre otras personas es limitada y desorganizada y, por lo tanto, existe un *abismo* entre lo que realmente podemos observar sobre ellos y lo que ellos mismos sienten, pretenden o creen. Cada uno de nosotros, por lo tanto, tiene que participar en uno u otro “desvío” intelectual para cerrar la brecha entre lo que se puede experimentar de inmediato sobre la otra persona y los estados psicológicos de esa persona (Asch, 1952).”

⁶⁵ Aquí, no debe confundirse la noción de empatía *des-centrada* con el planteamiento acerca de que en la interacción “no hay un abismo epistémico entre el yo el otro”. En particular, la ausencia de tal “abismo epistémico” refiere a que en el proceso de la interacción no existe un abismo entre lo que observamos del otro y lo que él mismo siente, pretende o cree, pues esto se nos revela directamente mediante sus acciones, y de esta manera conocemos su mente en la dinámica misma de la interacción a través del involucramiento emocional y del acoplamiento mutuo. Más adelante se verá esto con detalle.

(Schilbach, *et.al.*, 2013, p.4)

La diferencia de esta concepción con el enfoque de la 2ª persona, es que éste último afirma que en la interacción *experimentamos inmediatamente al otro* como sujeto, afirmación que se basa en la tradición fenomenológica (Asch, 1952; Heider y Simmel, 1944; Runeson y Frykholm, 1983; Schutz, 1972; Thinès et al., 1991), en donde se plantea que:

“La calidad de las acciones [de los otros] imbuje a las personas en una realidad viva. Cuando decimos que una persona tiene dolor, vemos su cuerpo como sentimiento de ello. No necesitamos 'imputar' conciencia a otros si percibimos directamente las cualidades de la conciencia en las cualidades de la acción. Una vez que vemos un acto que es hábil, torpe, alerta o imprudente, es superfluo ir "detrás" de él a [buscar] su sustrato consciente, porque *la conciencia se ha revelado en el acto*” (Asch 1952, p.158, las cursivas son mías).

De esta manera, el enfoque de 2ª persona resalta que *experimentar e interactuar directamente* con otros constituye nuestra principal manera de conocer sus mentes, pues sus cuerpos y acciones nos la develan inmediatamente. Por tanto, desde este punto de vista, no es necesario aplicar, de entrada, una rutina atribucionista elaborada para inferir sus estados mentales.

Pero, ¿hay evidencia empírica respecto a lo que se está diciendo? Sí la hay y a continuación la presento.

3.6.3. Evidencia empírica

- 2ª persona

Primero, ¿cuáles son las implicaciones de tomar una perspectiva epistémica de 2ª persona en la cognición social y qué evidencia hay al respecto?

La perspectiva de 2ª persona trata con un enfoque del conocimiento social basado en la interacción y en los involucramientos emocionales *entre* agentes, en donde el entendimiento de otras mentes está estrechamente relacionado con la experiencia y con los sentimientos de estar genuinamente involucrado, así como también, con las respuestas emocionales que tenemos hacia los demás, es decir, entender a los *otros* depende crucialmente del involucramiento emocional y de la capacidad de respuesta o receptividad a sus estados o acciones (Reddy, 2003, 2008; Rietveld, 2008). De esto resulta una visión de la cognición social que se da *entre* agentes (y no sólo ocurre en las representaciones de la mente individual), cuyo papel de *participantes* (y no de observadores) es que están

emocionalmente involucrados, mutuamente acoplados y co-creando significados en la interacción.

De manera relevante, el involucramiento emocional no sólo refiere a una forma de percibir los estados de los demás y de responder en consecuencia, sino que es una forma de involucrarse legítimamente con ellos, lo que contribuye a reconocerlos como *otros* y a comprenderlos, principalmente en términos de sensibilidad corporal, es decir, el involucramiento emocional nos permite ser sensibles al comportamiento socialmente relevante de los demás, lo que proporciona más posibilidades de acción e interacción entre los agentes.

A nivel neuronal, se ha sugerido una marcada diferencia fenomenológica entre las perspectivas de 2ª y de 3ª persona (2ª persona como *participante* y 3ª persona como *observador*), esto debido a que cada perspectiva involucra distintas regiones del cerebro y se correlaciona con diferentes patrones neuronales: la perspectiva de 2ª persona se correlaciona más con regiones involucradas en las emociones, motivaciones y evaluaciones y la de 3ª persona con regiones involucradas en procesos visuo-espaciales que favorecen el análisis y la explicación. Esta distinción ha inspirado un paradigma de investigación completamente nuevo en la neurociencia cognitiva social, a saber, la neurociencia de la 2ª persona.

Schilbach *et al.* (2006) muestran estos diferentes patrones neuronales que se activan en las dos condiciones mediante estudios de resonancia magnética funcional (fMRI). En concreto, los estudios exploraron la distinción que prevalece entre la experiencia de estar involucrado en la interacción social y la experiencia de ser un observador pasivo en la misma. En apoyo a ello, también se demostraron diferencias significativas en los patrones de activación que favorecen la percepción de expresiones faciales socialmente relevantes dependiendo de si estas expresiones eran dirigidas-a-uno o dirigidas-a-otros, lo que mostró que, en efecto, el punto de vista de un observador (*i.e.*, expresiones dirigidas-a-otros) recluta regiones del cerebro relacionadas con procesamientos visuoespaciales, mientras que la percepción de expresiones dirigidas-a-uno activa regiones del cerebro relacionadas con procesamientos emocionales y evaluativos.

En esta línea de investigación, también se ha sugerido que la perspectiva de 2ª persona impulsa las respuestas emocionales y corporales que desempeñan un papel crucial en la interacción, debido a que influyen en las redes de control de acción del cerebro y modulan

los procesos de integración sensoriomotora, lo que, a su vez, motiva el comportamiento y la acción; esto sesga y proporciona más posibilidades de interacción (Frijda 1986; Keltner y Haidt 1999; Schilbach et al., 2008).

En torno a ello, los autores han sugerido que las respuestas emocionales corporizadas pueden facilitar formas más elaboradas de la cognición social, como aquellas enfocadas en la predicción y la explicación (*i.e.*, como las rutinas atribucionistas), mientras que la ausencia de estas respuestas más básicas podría hacer más laboriosos esos procesos.

Ahora bien, en el campo de la psicología del desarrollo socio-cognitivo enfocado en la perspectiva de 2ª persona, se le ha otorgado un papel preponderante a la *interacción*. La sugerencia es que en la interacción ejercitamos nuestras habilidades sociales y nuestro conocimiento práctico social, puesto que es en donde conocemos directamente la mente del otro, participamos junto con el otro, nos modulamos con el otro y lo tratamos como otro. Por consiguiente, desde este punto de vista, se defiende que no hay un abismo epistémico entre el yo y el otro, pues sus pensamientos y sentires son revelados en la acción.

La razón es que, en la interacción social, más que ejercitar las capacidades reflexivas y deliberativas, primeramente ejercitamos nuestro conocimiento práctico (*know-how*) al tratar con otros como interactuantes y participando en conjunto en el habilitamiento de la interacción (Klin et al., 2003). Aquí, no sólo se trata de recolectar datos del otro en la interacción, sino que el conocimiento que tenemos de los demás es directo y reside en la dinámica misma de la interacción. En otras palabras, el conocimiento del otro surge de la interacción *entre* agentes y no sólo de la mente individual (De Jaegher et al., 2010). Entonces, desde este punto de vista, la dinámica de la interacción social contribuye y, en ocasiones, constituye nuestra conciencia de las otras mentes.

En esta línea, se ha investigado el desarrollo de la conciencia de las otras mentes mediante la perspectiva de 2ª persona, es decir, se ha estudiado cómo los infantes llegan a reconocer la existencia y la naturaleza de otras mentes mediante la interacción directa. A continuación, presento la evidencia empírica concerniente a ello.

Primero, la evidencia empírica muestra que los infantes, a los pocos minutos de nacidos, ya presentan un interés considerable por los rostros humanos y responden apropiadamente a las acciones faciales dirigidas a ellos, principalmente a acciones notorias como sacar la

lengua y abrir la boca (Kugiumutzakis 1998; Meltzoff y Moore, 1977; Nagy y Molnar 2004), lo que sugiere un sentimiento de conexión y cierto involucramiento emocional.

A los 2-5 días de nacidos, los infantes prefieren rostros con ojos mirándolos y presentan respuestas tempranas a las miradas dirigidas a ellos. La mirada mutua refleja el interés del infante por la mirada del otro y por el otro mismo, lo que sugiere un alto potencial de crear una situación de intensa reciprocidad (Farroni et al., 2002).

A los dos meses de edad, los infantes ya muestran una variedad de reacciones emocionales a la atención dirigida a ellos en la interacción, que van desde el placer (Wolff, 1987) hasta la angustia (Brazelton, 1986), así como también la ambivalencia de estos dos estados que se refleja en forma de sonrisas tímidas (Reddy, 2000), lo que sugiere un fuerte involucramiento emocional con el otro.

En esta línea, se ha mostrado que las complejidades del involucramiento emocional en las interacciones son sumamente importantes para el desarrollo del infante y, también, son el principal recurso de donde proviene el significado de uno mismo y de los demás, es decir, mediante el involucramiento emocional hacemos sentido de nosotros mismos y de los otros como otros (Fogel, 1993; Stern, 1985; Trevarthen, 1977; Cohn y Tronick 1989).

Al respecto, a los 2-3 meses de edad los infantes demuestran coordinación *intra* e *interpersonal* durante la comunicación cara a cara (Fogel 1993; Lavelli y Fogel 2002; Stern 1985; Trevarthen 1977), se muestran perturbados por la falta de coordinación temporal y afectiva (Cohn y Tronick 1989; Legerstee y Varghese 2001; Murray y Trevarthen 1985; Nadel y Tremblay-Leveay 1999), se ven afectados en su capacidad de llamar la atención por la historia de las respuestas afectivas maternas (McQuaid et al., 2009) y aprenden a preferir los niveles de contingencia interaccional con la que están previamente familiarizados (Bigelow y Rochat, 2006).

En este contexto, también se ha documentado que los involucramientos en 2ª persona promueven que el infante reconozca la *atención* de los otros (si ésta es dirigida a ellos, o a otra persona, o a algún objeto, etc.), lo que le permite desarrollar el reconocimiento de la atención en involucramientos triádicos, complejos y temporalmente extendidos (Reddy 2003, 2005, 2008). Concretamente, esto faculta a los infantes a la atención mutua y a la capacidad de responder al otro, lo que propicia la comunicación en contextos compartidos

que se desarrolla activamente en encuentros continuos (Clark 1996; Garrod y Pickering 2004).

Así, a partir de esta evidencia se ha argumentado que las experiencias de las otras mentes se experimentan primero y más intensamente en las interacciones de 2ª persona, en donde el agente que interactúa es abordado directamente por *otro* agente y reacciona a él en consecuencia, lo que contribuye significativamente al desarrollo de las competencias socio-cognitivas y a la comprensión social a lo largo de la vida. Además, las interacciones sociales así entendidas son facilitadas por la percepción directa de los estados mentales del otro (Rochat, 2003, 2015; Reddy, 2013).

En sentido concomitante, Vasudevi Reddy (2008) sostiene que los infantes que son demasiado pequeños para poseer el concepto de una “intención” o de una “creencia”, no simplemente perciben las acciones intencionales de los demás, sino que directamente imitan, responden, anticipan, invitan y se ajustan a las conductas ajenas, y antes de finalizar su primer año de vida, incluso, interrumpen deliberadamente (ver Reddy, 2008, p. 175). Con esto, la autora busca poner en primer plano las capacidades corporizadas que anteceden a los procesos cognitivos que defienden las teorías atribucionistas.

Asimismo, Reddy propone una poderosa liga entre la experiencia con el otro y una experiencia moral que “me obliga a responder”. En sus palabras:

“No sólo la experiencia de la otra persona es más inmediata y más poderosa en el involucramiento directo, sino que llama de ti una forma diferente de ser, una receptividad inmediata, un sentimiento en respuesta y una obligación de *responder* a los actos de la persona” (Reddy, 2008, p. 27).

De esta manera, en la psicología del desarrollo comienza a trazarse un vínculo directo entre la noción de *participante* y la noción de responsabilidad moral desde la perspectiva de 2ª persona, tema que ha sido sustanciado en el enactivismo por van Grunsven (2018). A esto vuelvo más adelante.

Reddy también explica cómo nuestro entendimiento del otro en involucramientos de 2ª persona es epistémicamente más rico: “sólo si estamos involucrados activamente con las personas podemos percibirlos como son” (Reddy, 2008, p. 29). Esto se debe a que “hay una circularidad [...] en términos de conocimiento de la mente: lo que sabemos de las mentes debe depender de nuestros involucramientos con ellas, pero estos involucramientos

deben depender en sí mismos de lo que sabemos de ellas. Cuanto más te relacionas con otras mentes, más debe haber con qué interactuar” (Reddy, 2008, pp. 31-32).

En esta línea, Reddy sugiere que, desde una perspectiva de 3ª persona de observación, explicación y predicción en una interacción, en realidad se nos escaparían aspectos que sólo pueden ser captados confiablemente desde una perspectiva de 2ª persona, las cuales nos permiten identificar y responder de manera confiable a las intenciones, deseos y necesidades de los demás.

Pero, concretamente ¿cómo llegamos a adquirir un sentido de nosotros mismos y de los demás precisamente como *otros*? Esta pregunta es relevante, no sólo para responder a los problemas computacionalistas, sino porque es un requisito para poseer genuinamente una perspectiva moral, tal como lo plantea Korsgaard. Veamos la respuesta a continuación.

- **Reconocer al otro como otro y significados compartidos**

Anteriormente presenté la evidencia empírica que sugiere que las complejidades del “involucramiento emocional” son el recurso principal para hacer sentido de nosotros mismos y de los otros como *otros*. En apoyo a ello, la psicología del desarrollo plantea que, de hecho, la capacidad de respuesta a la manifestación corporal de los demás como agentes expresivos intencionales ya está presente en los infantes humanos en edades muy tempranas, incluso, en recién nacidos. Revisemos la evidencia.

Primero, en relación con lo que se vio, Meltzoff y Moore (1977) han demostrado que los recién nacidos ya pueden imitar las expresiones faciales de su cuidador (ej. los movimientos de la lengua). Este mimetismo facial requiere que los recién nacidos tengan un sentido perceptivo de la conexión y de la diferencia entre ellos mismos y los demás, ya que deben ser capaces de “transponer los gestos y las expresiones faciales de los demás a su propia propiocepción y movimiento” (Fuchs y De Jaegher, 2009, p. 478). Esto sugiere además un sentimiento de conexión importante, pues no imitan pasivamente, sino que se involucran emocionalmente.

Segundo, recordemos que los bebés muestran de manera confiable una preferencia por los rostros humanos, a los que perciben y responden de una manera que generalmente no ocurre con los objetos inanimados. Esto desafía directamente la idea de que nuestra experiencia perceptiva de los demás es evaluativamente neutral y modalmente separada de la acción.

Ahora bien, aunado a estos puntos, se propone que las interacciones entre los infantes y sus cuidadores son “diálogos inter-corpóreos”, en donde el diálogo tiene un papel fundamental debido que es capaz de transformar tanto a los agentes como al significado y al reconocimiento del otro como tal. Como señala Reddy, “la comunicación y el diálogo crean y transforman al individuo y el ámbito del significado, pero también, curiosamente, parecen demostrar el reconocimiento de otro ser” (Reddy, 2008, p.66).

En apoyo a ello, la evidencia muestra que los infantes de 2 meses de edad ya son conversacionales e intencionalmente comunicativos, lo que se ve mediante “las sonrisas, los arrullos y la proto-charla del bebé” (Reddy 2008, siguiendo a Bråten y Trevarthen, 2007), lo que sugiere, no sólo el reconocimiento del otro, sino también el acoplamiento a la interacción y las intenciones comunicativas del infante para con los demás (esto lo han defendido numerosos psicólogos del desarrollo como: Colwyn Trevarthen; Mary Catherine Bateson; Lynne Murray; Daniel Stern; entre otros).

Asimismo, se ha argumentado que existe un reconocimiento mutuo y una capacidad de respuesta a la expresividad emocional y a la intencionalidad del otro mediante las interacciones en las que se negocia la “toma de turnos” (Jaffe et al. 2001; Goldstein y Schwade, 2008). Crucialmente, estas interacciones muestran la percepción adecuada de la estructura de la interacción social por parte del infante, lo que contribuye a que la interacción en sí misma sea considerada como un vehículo para la adquisición de conocimientos, por ejemplo, para la guía del desarrollo vocal y del aprendizaje del lenguaje (Bruner 1983; Goldstein y Schwade 2010), tanto en términos de percepción del habla (Kuhl 2007; Kuhl et al. 2003) como en la toma de turnos (Goldstein y Schwade 2008; Goldstein et al. 2009; Jaffe et al. 2001).

En sentido concomitante, los adultos también muestran el reconocimiento mutuo y la capacidad de respuesta a la expresividad y a la intencionalidad del otro en sus interacciones, pues continuamente coordinan turnos, se acercan a los acentos de los demás (Giles et al., 1992), balancean su cuerpo en sincronía con el otro (Shockley et al., 2003) y coordinan visualmente su atención a través de movimientos oculares sincronizados (Richardson et al., 2007).

Para generar más apoyo a esta afirmación acerca del reconocimiento mutuo y la receptividad, se ha llevado a cabo una serie de experimentos llamados “experimentos de perturbación”; los cuales son realizados a infantes desde los 3 meses de edad. Lo que han

mostrado estos experimentos es que, ante las fallas en la interacción, los infantes responden con angustia, lo que sugiere que ellos tienen expectativas de las acciones de la otra persona en relación con ellos (Reddy, 2008, p. 74, siguiendo a Ed Tronick).

El argumento es que los infantes son perceptualmente abiertos y están en constante sintonía con la manifestación corporal expresiva de los demás. No obstante, considerar a los demás como seres intencionales expresivos no es algo automático, sino que es negociado constantemente en la interacción. Así, lo que se adquiere a través de la interacción es un conocimiento pre-reflexivo de cómo tratar con los demás: cómo compartir el placer, llamar la atención, evitar el rechazo, restablecer el contacto, etc. De esta manera, “el infante adquiere esquemas interactivos específicos [...] que son necesarios para mantener la respectiva interacción” (Fuchs y De Jaegher, 2009, p. 481).

El psicólogo del desarrollo Ed Tronick ha denominado “expansión diádica” a este creciente desarrollo y dominio de las habilidades de la interacción efectuadas mediante involucramientos en 2ª persona⁶⁶.

Entonces, la interacción empática diádica es motivada por la continua sintonía afectiva y por involucramiento emocional, pero, a su vez, por los desajustes y las perturbaciones también.

En esta línea experimental, también se ha sugerido que los infantes a los 4 meses de edad ya cooperan para ser cargados (Reddy, 2008, en Grunsven, 2015, p.237), lo que logran por adaptarse y anticiparse corporalmente a las posturas habituales que usan sus cuidadores para levantarlos, *i.e.* se modulan a los actos intencionales dirigidos a ellos.

Asimismo, antes del año de edad (9 meses) se burlan de otros, es decir, los infantes comienzan a buscar activamente y a provocar las perturbaciones en la interacción mediante actos de burla (Reddy, 2008), lo que refleja entendimiento intencional del otro, ajustes del comportamiento propio a los significados emergentes de la situación y a la percepción de la estructura interaccional. Como señala Reddy, solo podemos *burlarnos* de los demás si

⁶⁶ Reddy captura la propuesta de Tronick al decir que: “Las experiencias del infante pueden expandirse en complejidad dentro de una base diádica: el involucramiento diádico lleva al infante un paso más allá de sus capacidades, atrayéndolo hacia habilidades e incrustándolo en patrones de sentir y hacer en los que de otro modo no habría tenido acceso. Estos patrones brindan coherencia a la experiencia infantil, o incluso adulta, y no solo reflejan el conocimiento actual, sino que también influyen en lo que podemos aprender y saber. Ciertamente, sirven para invitar a los infantes a nuestra cultura” (Reddy, 2008, p. 81).

jugamos con sus intenciones sobre la base de una sensibilidad a la estructura de una interacción ya establecida.

Así, pues, el planteamiento puntual es que los significados se negocian de manera interactiva y las vidas expresivas e intencionales de los demás están perceptualmente disponibles, de tal suerte que hacemos sentido del otro como *otro* en la interacción, lo que cuenta con apoyo empírico sólido y tiene sentido desde una perspectiva del desarrollo (Reddy, 2008, p.155).

No obstante, aún falta responder de manera más precisa ¿cómo es que el agente hace sentido de sí mismo? o bien ¿cómo es que el agente logra auto-individuarse? A continuación, presento los argumentos concernientes a ello.

- **Auto-individuación**

Se arguye que la empatía y el involucramiento emocional en la interacción, con sus momentos constitutivos de ruptura, perturbación y de desajustes, contribuyen significativamente a crear una nueva relación experiencial con los otros y con el entorno desde una perspectiva des-centrada. Esta perspectiva des-centrada del mundo da lugar a la habilidad de auto-monitorear la conducta y la cognición de uno mismo, de lo que emerge una nueva auto-relación corporizada empática con el otro y con el entorno, capacidad con la que se auto-individua el agente mismo (Grunsvén, 2015, p.246). Veamos esto con más detalle.

La primera razón es que las cosas en el mundo las percibimos de manera distinta en comparación con cómo percibimos y experimentamos nuestro propio cuerpo vivo con el que interactuamos en el mundo; se arguye que de éste emerge una subjetividad corpórea que constituye un *único* punto de vista.

El planteamiento es que las cosas en el mundo no las percibimos únicamente en términos de que están *dadas* aquí y ahora, sino también en términos de su potencial accesibilidad desde una multiplicidad de puntos de vista que permiten diversas posibilidades de intercambios dinámicos con el entorno, lo que no ocurre con el propio cuerpo vivo, puesto que éste, desde la perspectiva de 1ª persona, siempre está “aquí” y no “allí” (siempre está en el *punto-cero*) (Thompson, 2007, p. 392), de tal forma que nuestra subjetividad está ligada a un sentido completo de nuestro cuerpo en su propia orientación espacio-temporal que constituye un *único* punto de vista con el que nos *abrimos* al mundo.

Esto constituye una diferencia fenomenológica entre cómo percibimos los objetos del mundo y cómo nos *experimentamos* nosotros mismos como un *yo* con un cuerpo vivo propio con el que interactuamos con el mundo, de lo que emerge una nueva auto-relación corporal que permite explicar la perspectiva des-centrada que tenemos los humanos con el entorno.

En particular, esta experiencia de auto-individuación da lugar a la posibilidad de des-centrar la propia perspectiva y en su lugar, en el caso social, experimentar cómo somos percibidos por los demás, principalmente mediante la empatía, lo que contribuye a que surja la habilidad para auto-monitorear la conducta y la cognición de uno mismo.

Me explico, el pensamiento de la auto-individuación se ha asociado con la cognición social y a partir de ello se plantea que es mediante la empatía, específicamente mediante *la experiencia de uno mismo como un otro para el otro*, que uno obtiene una visión del propio ser corpóreo más allá de la perspectiva de la primera persona (ver Thompson, 2007, p. 392, las cursivas son mías). De ello, Zahavi ha argumentado que es “la internalización de la visión del otro sobre uno mismo [lo que] da lugar a la capacidad de autocontrolar críticamente el propio comportamiento y la propia cognición” (Zahavi, 2014, p. 230); característica que contribuye significativamente a la cognición social desde un enfoque corporizado.

Para hacer más clara esta idea acerca de cómo logramos auto-individuarnos, recorro a la noción de “*subjetividad sensoriomotora*” (Maiese, 2011), que es un modo primitivo de consciencia corpórea que provee el fundamento para un sentido del “yo”. Primero, los organismos obtienen el sentido más básico del “yo”, que es un sentido corpóreo del “yo” que surge de procesos metabólicos (el *yo-corpóreo*), siendo éste, un “yo” a nivel de metabolismo celular, pues implica la producción y manutención de una identidad dinámica ante el cambio material del entorno (Maiese, 2011, p. 105).

Segundo, el *yo-corpóreo* da lugar a un *yo-sensoriomotor*. El *yo-sensoriomotor* es producido por los procesos sensoriomotores; es un “yo” a nivel de percepción y acción, contorneado por nuestra corporalidad esencial, es decir, nuestro cuerpo define sus contornos, lo que otorga un carácter muy definido a nuestra experiencia perceptual, estableciendo un centro

y una periferia; así que, conforme a la interacción con el mundo, uno desarrolla una consciencia implícita de la distinción entre el "yo" y el "otro", o bien, entre el "yo" y lo "otro".⁶⁷

El argumento es que la subjetividad sensoriomotora (dicho modo primitivo de consciencia corpórea) está ligada a un sentido completo de nuestro cuerpo en su propia orientación espacio-temporal que conforma un *único* punto de vista. Así, nuestras dinámicas corpóreas informan nuestro conocimiento y comprensión del mundo, sirven como el recurso de la estructura espacio-temporal de la experiencia consciente y nos permiten interactuar con el mundo desde un único punto de vista.

De esta manera, tal consciencia primitiva corpórea forma nuestro entendimiento temprano del mundo en percepción y acción, así como también nuestro sentido del "yo", de tal suerte que la subjetividad sensoriomotora constituye las bases para el contenido sensorial asociado con el sentido de percepción, pero también para las formas más sofisticadas de cognición incluyendo actos de pensamiento, pues presuponen subjetividad sensoriomotora. Esta afirmación implicaría que los estados mentales involucran este tipo de consciencia.

Ahora bien, Reddy corrobora estos planteamientos enactivos y sugiere que esta nueva relación corporal con uno mismo y la forma en que nuestro cuerpo está "marcado" como una "entidad separada", surge en la primera infancia a través de la interacción, en donde intervienen fundamentalmente los actos de atención cargados de emociones. Esto significa que la adquisición de esta nueva relación corporal con uno mismo es inseparable de la experiencia del propio cuerpo como algo digno de ser atendido y abordado en involucramientos emocionales para la interacción, lo que favorece una perspectiva descentrada con el mundo (natural y social) (ver Reddy, 2008; Grunsvén, 2015).

Así, pues, hasta aquí he presentado la evidencia que apuntala las predicciones de la cognición social enactiva, mismas que potencializan el campo moral, por lo que ahora corresponde revisar la evidencia que apoya las facultades morales corporizadas.

3.6.4. Evidencia para facultades morales tempranas: conducta cooperativa mínima

Como primer paso, analizo la "intencionalidad compartida" o "intencionalidad colectiva", la cual es aquella en la que la colaboración y la ayuda surgen de la estructura de la actividad

⁶⁷ Aunque el "yo" en sí mismo es esencialmente corporizado y está enraizado en nuestra naturaleza biológica, no es reducible a ella.

compartida sin necesidad de postular ningún factor más, como una tendencia altruista innata para apoyarse mutuamente (Satne, 2018, p.13). La intencionalidad colectiva es la capacidad de participar en un grupo y de acuerdo con este punto de vista, los individuos muestran intencionalidad colectiva cuando forman parte de un grupo que se propone hacer algo, lo que denota su aspecto *relacional*.

La hipótesis es que la capacidad de adaptarse e interactuar con otros para alcanzar objetivos compartidos es anterior a la intencionalidad conjunta que implica representar estados mentales y las perspectivas de otros, así como también, que la colaboración se basa en un sentido de involucramiento y pertenencia a un grupo que comienza a desarrollarse muy temprano en la infancia. (Hobson, 2002; Reddy, 2015; Over y Carpenter, 2009, Buttelmann et al., 2012).

Así, se sugiere que hay un tipo de intencionalidad colectiva *mínima* reflejada como un comportamiento grupal que se integra con formas básicas de acción conjunta, lo que ayudaría a desarrollar cogniciones más sofisticadas, tales como aquellas que involucran representaciones mentales y toma de perspectiva, entre otras.

La intencionalidad colectiva *mínima* designa formas especiales de acción y de afiliación grupal que se despliegan en la acción conjunta, mismas que surgen de las normas, prácticas y hábitos compartidos más que de representaciones mentales (Satne, 2018, p.15). Los autores señalan que tal variedad mínima de intencionalidad colectiva se ve reflejada en las interacciones diádicas predominantes en los primeros años de vida.

Ahora bien, lo que caracteriza los comportamientos de intencionalidad compartida *mínima* son formas de interacción en las que el papel de cada participante está determinado y estructurado por la actividad compartida. En edades muy tempranas, esto se puede reflejar por la capacidad que tienen los infantes de coordinar y sintonizar sus acciones con las acciones y reacciones dirigidas a ellos en actividades muy simples, tales como acomodarse corporalmente para ser levantados o para ser cambiados.

Al respecto, se ha mostrado que los bebés de 2 a 4 meses de edad se adaptan y anticipan corporalmente a las posturas habituales que usan sus cuidadores para recogerlos, por ejemplo, de los asientos del auto (ver Reddy, 2015; Reddy et al., 2013); y, los infantes de 6

meses lo hacen, pero ahora para cumplir físicamente con las directivas intencionales (incluidas las lingüísticas, Reddy, 2012, 2015; Rochat, 2015)⁶⁸.

Crucialmente, estos acoplamientos en las interacciones tempranas no requieren que los infantes planifiquen con anticipación o predigan y representen intelectualmente los estados mentales de los otros, sino que, más bien, se trata de *acoplamientos* que responden a la sintonía inmediata de la participación diádica.

Pero, concretamente ¿cuáles son los mecanismos psicológicos que subyacen a la intencionalidad colectiva mínima?

El punto de partida es reconocer que los infantes en edades muy tempranas tienen disposiciones para involucrarse en un tipo de comportamiento grupal que no requiere que representen las perspectivas del otro sobre sus acciones, sino que surge esencialmente de la sensibilidad a la pertenencia del grupo, de una perspectiva socio-céntrica (más que individual) y del acoplamiento a la estructura de la interacción que está permeada de prácticas compartidas.

Bajo este supuesto, Satne (2018) identifica tres dimensiones psicológicas que considera suficientes para dar sentido a la intencionalidad colectiva mínima: 1) el conformismo social; 2) formas básicas de coordinación interpersonal espontánea; y 3) formas no sofisticadas de cognición social. Los tres elementos son cruciales para dar sentido a la acción compartida y al aprendizaje social, lo que conduce al surgimiento y desarrollo de comportamientos colectivos y sociales más complejos.

El “conformismo social” describe un conjunto de mecanismos capaces de moldear las disposiciones de los niños para que coincidan con el comportamiento de otros miembros de la comunidad y además para continuar monitoreándolos para su cumplimiento (Haugeland 1998). Estos mecanismos de conformidad social hacen que surja la práctica de aprender y enseñar y no requieren que las personas cumplan deliberadamente con las reglas (en principio), sino desde acoplamientos más básicos. En lugar de representar reglas y evaluaciones de los demás sobre sus acciones, las personas son sensibles a las

⁶⁸ Cabe resaltar que en la evidencia computacionalista, esto ocurre entre los 8 y 16 meses de edad (Hamlin, 2015).

evaluaciones de los demás a través de vías cognitivas más mínimas, como la sintonización emocional con la aprobación y desaprobación de los demás (ver Satne 2018, p.18).

Esta capacidad está bien documentada y se ha estudiado experimentalmente bajo distintos nombres: por ejemplo, mecanismos para “moldear la mente” (Zawidzki, 2012), para la “adaptación mutua” (Kiverstein y Roepstroff, 2016), para la “regulación normativa psicológica popular” (McGeer 2007; 2015), “imitación de prejuicios del grupo” (Buttelmann et al., 2013), “pedagogía natural” (Csibra y Gergely, 2009), “cumplimiento normativo” (Rakoczy, Warneken y Tomasello, 2008) y para la imitación/emulación (Tomasello, 1996; 1999; Call et al., 2005; Tomasello y Rakoczy, 2003; Tennie, Call, y Tomasello, 2009; 2012).

La “coordinación”, por su parte, refiere a la capacidad que tienen los individuos para seguir a otros y para ajustarse al comportamiento del grupo (Pacherie, 2013; 2015). En torno a ello, existen estudios que muestran que hay una tendencia innata en los humanos para coordinar sus propios movimientos corporales con los de los demás (por ejemplo, al caminar), y, progresivamente, con actividades más sofisticadas (ver Knoblich et al., 2011).

Asimismo, se han estudiado diversas “dinámicas interactivas” que revelan la presencia de mecanismos involucrados en la coordinación espontánea de ajustarnos a los demás y a la estructura de la interacción (ver De Jaegher, Di Paolo y Gallagher, 2007).

Esta capacidad es crucial para comprender cómo las actividades coordinadas conjuntas son posibles en edades muy tempranas, así como también, para comprender su desarrollo en actividades coordinadas más complejas en etapas posteriores, cuando éstas se combinan con otras habilidades más sofisticadas, incluidas las que implican planificaciones complejas.

Por último, las “formas básicas de la cognición social” son fundamentales para entender la capacidad que tienen los infantes para involucrarse en interacciones de intencionalidad colectiva *mínima*, puesto que ayudan a explicar cómo los niños pueden identificar la aprobación o desaprobación de los demás al sintonizarse con las reacciones de otras personas.

Primero, se reconoce que al menos algunos estados mentales (como las emociones o intenciones -Gallagher, 2008-) son directamente observables. Como se vio anteriormente, hay estudios en psicología del desarrollo que han proporcionado evidencia de que los encuentros cara a cara (2ª persona) implican una comprensión perceptiva directa de las

emociones de los demás y de sus acciones dirigidas intencionalmente, incluso en edades muy tempranas (Reddy et al., 2013; Rochat, 2003, 2015). Con ello, se ha argumentado que en tales encuentros no es necesario implicar el uso explícito o implícito de representaciones, inferencias o simulaciones (Gallagher 2001, Zahavi 2008, Hutto 2014, Reddy et al 2013, Zahavi 2014, Satne 2014).

Bajo esta evidencia, los autores plantean que las estrategias atribucionistas (TT y/o TS) no serían requeridas para que surjan formas primitivas y mínimas de la cognición social, lo que permite explicar los involucramientos de intencionalidad compartida *mínima* en infantes de 2 meses de edad (Reddy, 2008; Rochat y Striano, 1999).

Segundo, es preciso tener en cuenta que la intencionalidad compartida se fundamenta en un sentido de participación y pertenencia al grupo que comienza a desarrollarse muy temprano en la infancia (Hobson 2002; Reddy 2015). Asimismo, se plantea que la capacidad de adaptarse al comportamiento de grupos o de otro individuo es lo que permite que los niños comiencen a participar en actividades compartidas cada vez más complejas con los demás.

En este contexto, lo que se ha buscado es demostrar que las conductas de ayuda están moduladas, al menos en parte, por la participación y la pertenencia al grupo desde edades muy tempranas. La evidencia respecto a niños mayores sigue siendo controversial, no obstante, hay experimentos con bebés muy pequeños, de 2 a 8 meses de edad, que brindan evidencia que apuntala la intencionalidad colectiva *mínima*.

Como ya se ha visto, Reddy (2003) argumenta que los infantes, a partir de los 2 meses de edad y antes de cualquier evidencia de atención compartida, responden a los demás en interacciones ricas en emociones, así como también, que estas interacciones tempranas les brinda un conocimiento práctico de cómo la otra persona se dirige intencionalmente hacia ellos, como lo muestra la evidencia que sugiere que los bebés de esta edad anticipan acciones intencionales dirigidas a ellos, por ejemplo para ser levantados o cambiados. Y, a los 6 meses de edad ya son capaces de cumplir instrucciones intencionales, incluyendo las lingüísticas (Reddy 2012, 2008; Rochat, 2015).

Esto proporciona evidencia para la idea de que la (inter)acción, que implica respuestas intencionales y reacciones estructuradas complementariamente con los demás, es anterior

a comprender y a distinguir dos perspectivas individuales distintas para la atribución de estados mentales y para la complementación de acciones conjuntas.

También hay estudios realizados a infantes de 8 meses de edad que apoyan la perspectiva colectiva. Ramenzoni y Liszkowski (2016) efectuaron una serie de experimentos en los que bebés de 8 meses, que aún no se comunican mediante el gesto de señalar, utilizan acciones instrumentales con la aparente expectativa de que el otro adopte y prediga sus objetivos.

El experimento pone a prueba las acciones de infantes de 8 meses de edad para alcanzar objetos en dos escenarios contrastantes: cuando un adulto está presente y cuando nadie está presente, cambiando selectivamente la posición de los objetos deseados con respecto a los límites de acción de los bebés. Los autores descubrieron que los infantes, en el escenario individual, solo alcanzaban el objeto cuando estaba dentro de sus límites de acción, mientras que selectivamente aumentaron su alcance de objetos lejanos fuera de sus límites de acción cuando un padre o un adulto menos familiar se sentaba a su lado.

Desde la perspectiva colectiva, estos resultados son interpretados de la siguiente manera:

“[L]os infantes esperan que el adulto contribuya a sus acciones porque aceptan que el adulto se involucre en una acción colectiva con ellos, no porque tengan una comprensión intelectual de la noción de intención aplicada a ellos mismos y al adulto. Más bien, el bebé tiene una *comprensión práctica* de la actividad en la que está involucrado como una actividad conjunta en la que el adulto también tiene un papel que desempeñar, a saber, en este caso, el de proporcionarle el objeto que está tratando de alcanzar” (Satne, 2018, p.25).

Así, en esta línea empírica, lo que se ha buscado es mostrar que las conductas de ayuda están moduladas, al menos en parte, por la participación y la pertenencia al grupo desde edades muy tempranas, en donde la acción conjunta no requiere necesariamente representar los estados mentales ajenos (en principio), sino que ésta surge de las normas y de los hábitos de fondo (*background*) compartidos que, a su vez, proporcionan una plataforma para el surgimiento de capacidades de representación, lo que conduce al surgimiento y desarrollo de comportamientos colectivos y sociales más complejos y sofisticados (Satne, 2018).

Ahora bien, todas estas dinámicas sociales son sugeridas como esenciales para la configuración del dominio moral, de tal suerte que los escenarios de intencionalidad compartida serían el resultado de los nichos cooperativos en donde el infante está

incrustado desde su nacimiento; característica que contribuye decisivamente tanto al desarrollo moral como a la unicidad cognitiva del humano.

De ello, el enfoque enactivo sostiene que los juicios morales y el comportamiento moral serían corporizados en naturaleza, habilitados por procesos emocionales y automáticos, y que siempre están incrustados en contextos socio-culturales particulares. De esta manera, la moralidad se basaría parcialmente en hábitos corporizados que se han cultivado en contextos concretos de interacciones sociales (Varela, 1999).

En este punto es conveniente aclarar que, como es de esperarse, las metodologías empleadas por cada marco conceptual y sus recursos habilitadores son distintos, por lo que sus interpretaciones empíricas divergen considerablemente. Como se ha hecho ver, la metodología que apuntala las afirmaciones enactivas parten de la necesidad de efectuar investigaciones en escenarios genuinamente interactivos, es decir, la finalidad, desde este punto de vista, es realizar investigaciones que permitan el estudio de encuentros sociales en tiempo real de una manera verdaderamente interactiva y relacional.

Así, pues, si se toma en serio la perspectiva relacional, dinámica y colectiva del enactivismo tenemos buenas razones para pensar que las interacciones sociales y las prácticas compartidas no solo contribuyen a *gatillar* nuestras facultades morales, sino que realmente son uno de sus recursos constitutivos. Como señalan Di Paolo et al. (2018): “Las interacciones sociales, y las prácticas que sustentan, son el horno donde se forjan tanto la cultura como la cognición; son una fuente, no un punto de encuentro” (Di Paolo et al., 2018, p. 133).

Por último, cierro este capítulo respondiendo enfáticamente a la pregunta que planteé en el inicio de esta sección, según mi interpretación: ¿qué resuelve empíricamente el enactivismo que el computacionalismo no? las respuestas las trazo a manera de consecuencias:

3.6.5. Consecuencias

Tal como se vio, las contribuciones enactivas a la cognición socio-moral (que derivan naturalmente de su metodología relacional) son las siguientes:

1. Explica cómo hacemos sentido del otro como *otro*, lo que no es trivial, sino que es una característica constituyente de la moral.
2. Explica cómo nos auto-individuamos como agentes autónomos.

3. A partir de (1) y (2), se obtiene una perspectiva des-centrada del mundo que forma parte de nuestra perspectiva moral.
4. Explica la intencionalidad compartida (*mínima*) y sus mecanismos subyacentes, lo que contribuye, al menos en parte, a formar una plataforma constituyente del desarrollo moral, en tanto que da lugar a formas sociales más complejas como la moral.

Estas características, a su vez, contribuyen significativamente a tres aspectos de la moral como estructura social: 1) permiten develar su aspecto situado mediante los significados compartidos, 2) explican la co-constitución agente-agente (lo que da lugar a la dependencia recíproca), y, 3) posibilitan explicar el papel de un agente *activo* genuinamente empotrado en un contexto permeado de normatividades y prácticas compartidas. Con esto en mente, doy paso al siguiente capítulo en el que condenso las aportaciones computacionalistas y enactivas, lo que me permitirá habilitar concretamente mi propuesta general. Previo a ello, resalto las contribuciones que brinda el enactivismo en el desarrollo moral:

Cronología (intervenciones corporizadas):

Recién nacidos- muestran sentimientos de conexión e imitan de una manera no-pasiva, sino que se involucran emocionalmente, lo que ya sugiere empatía.

2 semanas- contagio de llanto como un tipo de empatía.

2 meses- inician comunicación y toma-de-turno, ya son receptivos al involucramiento emocional y a los estados emocionales de su cuidador y con ello comienzan a mostrar conductas de cooperación *mínima*.

2-3 meses- muestran coordinación intra e interpersonal durante la comunicación cara a cara, denotan atención mutua y la capacidad de responder a otro.

2-4 meses- reconocimiento del otro, receptividad, acoplamiento a la estructura de la interacción y cooperación *mínima* catalizada por la sensibilidad a la pertenencia de grupo.

En los primeros 6 meses- sensibilidad a la pertenencia de grupo y sensibilidad al sufrimiento de otros. (Cabe decir que en el computacionalismo se pensaba el surgimiento de la noción *socio-céntrica* hasta los 3 años de edad, mientras que, en el enactivismo, desde los 2 meses de edad).

6 meses- cumplen con las directrices, incluso las lingüísticas.

8 meses- muestran comprensión práctica de la actividad conjunta.

9 meses- atención e intencionalidad compartida, involucramiento en situaciones triádicas, sintonía afectiva que se complejiza, entendimiento intencional del otro, ajustes del comportamiento propio a los significados emergentes de la situación y a la percepción de la estructura interaccional.

18 meses preocupación empática y la sensibilidad a la pertenencia de grupo se relaciona fuertemente con conductas pro-sociales.

Capítulo 4

Hacia una propuesta integrativa computacionalista-enactiva de la cognición socio-moral y sus implicaciones normativas

4.1. Introducción

En los capítulos anteriores revisé los postulados del computacionalismo y del enactivismo en torno a la cognición socio-moral humana, así como también sus implicaciones normativas. Al mismo tiempo, abordé algunas de las dificultades centrales que cada uno de ellos presenta, lo que me permitió mostrar cuáles son las virtudes y los problemas de ambas teorías.

En particular, se vio que tanto el computacionalismo como el enactivismo presentan algunos problemas respecto a los cinco temas en los que dividí la presente tesis: 1) Evolución, 2) Cognición, 3) Cognición social, 4) Cognición moral y 5) Normatividad. A fin de seguir el hilo conductor, a continuación, los menciono brevemente, comenzando con los computacionalistas y posteriormente con los enactivos.

Concerniente al relato evolutivo, el computacionalismo presenta algunos problemas al momento de cuestionar los principios de la escuela de Santa Bárbara, debido a que apela a la selección natural como principal fuerza evolutiva, lo que no es suficiente para dar cuenta de la moral, o bien, para explicar el tipo de conocimiento moral general que nos es dotado evolutivamente; pues para ello, he sostenido que es necesario recurrir al entrelazamiento de los procesos evolutivos biológicos con los procesamientos culturales.

Referente al modelo cognitivo computacionalista, se vio que éste implica una separación modal entre la percepción y la acción, lo que deriva en el supuesto de que la percepción es *neutra* en tanto que la percepción representa características evaluativamente neutrales del mundo. Este supuesto es particularmente problemático al momento de abordar lo correspondiente a la cognición social, puesto que no explicaría uno de sus *explanadums* centrales, a saber, entender a los otros como *otros* (Reddy, 2008).

En esta línea, entonces, la rutina atribucionista de la cognición social computacionalista presenta al menos dos problemas: 1) debido a que concibe la percepción como *neutra*, no logra explicar cómo hacemos sentido de nosotros mismos y de los otros como *otros*, y; 2)

concibe que el significado de la acción humana se encuentra en los estados mentales, los cuales son inobservables y por tanto inaccesibles a la percepción.

Respecto a la cognición moral, se vio que si bien, la explicación de los procesos sub-personales implicados en ella son necesarios, también propuse que se requiere una metodología interaccionista para dar cuenta de la misma. Esto debido a que, en mi interpretación, es necesario otorgar autonomía a los procesos supra-personales desde su aspecto relacional, lo que considero que debe ganar un lugar en el estudio de la moral.

Por último, concerniente a las implicaciones normativas, he defendido que lo anterior implica estudiar la moral desde una metodología relacional y como estructura social, lo que contribuye al menos a tres aspectos de la normatividad moral: 1) permite develar su aspecto situado mediante los significados compartidos que emergen momento a momento de la interacción en marcha, 2) explica la co-constitución agente-agente (lo que da lugar a la dependencia recíproca), y, 3) posibilita explicar el papel de un agente *activo* genuinamente empotrado en un contexto permeado de normatividades y prácticas compartidas.

Ahora bien, el enactivismo, por su parte, también presenta dificultades al postularse como estrategia única y excluyente, pues como se vio, he propuesto que es necesario recurrir a los procesamientos cognitivos representacionales para dar cuenta de la cognición socio-moral humana en toda regla (Bohl y van den Bos, 2012), en particular, a los procesamientos de información guiados por principios, pues el aspecto relacional del nivel supraindividual no es suficiente para explicarla.

En este contexto, se hizo ver que el enactivismo presenta dificultades en lo siguiente: respecto al argumento evolutivo, el enactivismo se suscribe plausiblemente a la hipótesis de la construcción de nicho, que, si bien explica el potencial de los procesos culturales en las explicaciones evolucionistas de la moral, también es cierto que está fundamentalmente anclada a la evolución biológica, por lo que el papel de la selección natural y del *adaptacionismo* es esencial y como tal debe ser resaltado y enfatizado⁶⁹.

Concerniente a los principios de la cognición enactiva, de nuevo, se vio que además de la metodología relacional, se requiere también de una metodología computacionalista para dar cuenta del funcionamiento interno del organismo en términos de qué sucede en la

⁶⁹ Esto no significa que el enactivismo lo niegue, sino que en mi propuesta he sugerido otorgarle un papel más contundente para sustentar mi hipótesis general.

estructura interna del agente para explicar las formas en que éste procesa e ingresa la información. Concretamente, arguyo que es necesario conocer los procesamientos de información cargados de contenido y guiados por principios que constituyen la cognición.

Respecto a la cognición social y acorde con lo anterior, en mi interpretación, es necesario recurrir a la perspectiva epistemológica de 3ª persona (ToM), al uso de reglas y a una arquitectura cognitiva altamente interactiva que sea capaz de explicar los procesamientos funcionalistas-cognitivos del sujeto y de involucrar de manera sustancial la información pragmática, *i.e.* he recurrido fundamentalmente a la arquitectura cognitiva súper-modular, tal como se definió.

Respecto a la cognición moral, se hizo ver, mediante la evidencia empírica, que es muy amplio y complejo el espectro moral, de tal forma que requerimos de los postulados de dicha súper-modularidad para dar cuenta de ella satisfactoriamente, lo que permitirá contar con diversas estrategias disponibles para efectuar las tareas relevantes.

Por último, concerniente a las implicaciones normativas, sostengo que además de las aportaciones enactivas, también es necesario recurrir al argumento del conocimiento moral general evolutivamente dotado para asegurar la objetividad moral y evitar con ello la apropiación cultural de la misma, lo que es otorgado plausiblemente por el computacionalismo.

Así, pues, pese a las dificultades que deja cada marco conceptual, he considerado que los mismos son los que presentan una oportunidad para proponer el posible marco integrativo que planteo, en donde me interesa mostrar que algunas versiones de ambas teorías son necesarias, por motivos distintos, para comprender plenamente la cognición socio-moral humana.

Para ello, sería conveniente dejar de evaluar al computacionalismo y al enactivismo como teorías rígidamente opuestas y en su lugar considerar que hay un espacio apropiado para cada una de las capacidades y procesamientos funcionalistas y corporizados respectivamente, y de ello, observar cómo podrían complementarse adecuadamente y aceptar que ambas pueden ser parte de una misma explicación, aunque se trate de elementos independientes entre sí.

Enfáticamente, desde un pluralismo pragmático, se trata de analizar los dos modos cognitivos y los aspectos que cada uno captura: el computacionalismo enfocado en los

procesamientos de información guiados por principios y habilitados por representaciones, y, el enactivismo, enfocado en las dimensiones fenomenológicas, afectivas, interactivas y corporizadas desde su propia concepción *relacional* de agencia e interacción.

Ahora bien, para abrir la posibilidad de tal integración computacionalista-enactiva, recurro a la arquitectura cognitiva llamada *programa informático de inteligencia artificial* (PI en adelante), la cual, sostengo, acepta dos niveles de interpretación (el funcionalista y el corporizado). Propongo que esta arquitectura permite que un nivel ejecute los procesamientos computacionalistas y el otro los procesamientos enactivos, ya sea en un modo de cooperación, o bien, de modo independiente.

Para dar pie a mi propuesta, es pertinente señalar las condiciones mínimas que deben cumplirse para que un caso cognitivo cuente como un caso computacionalista o como un caso enactivo, las cuales derivan de las implicaciones de cada marco conceptual. Esto lo efectúo con la finalidad de mostrar que mi propuesta de integración incluye dichos elementos de ambos.

Asimismo, se deben hacer explícitos los compromisos asumidos para lograr sostener tal integración. Por lo que dedico los siguientes dos sub-apartados a la exposición de estos dos puntos, para después presentar el desarrollo de la propuesta misma.

4.1.1. Condiciones mínimas para que sea un caso computacionalistas o sea un caso enactivo

Condiciones mínimas para que sea un caso computacionalista

- La presencia de procesamientos guiados por principios y habilitados por representaciones.

Observación: La *representación* es entendida en un sentido mínimo como una estructura portadora de información.

Condiciones mínimas para que sea un caso enactivo

- La presencia de (inter-)acciones con el entorno habilitadas por bucles de percepción-acción que permiten más posibilidades de actuación y de significación para el agente.

Observación: La *acción* es entendida mínimamente como acoplamientos dinámicos agente-entorno (o agente-agente) en un confrontamiento inmediato con el mundo.

4.1.2. Compromisos asumidos que habilitan la propuesta integrativa computacionalista-enactiva

Para dar pie al marco integrativo que propongo asumo los siguientes dos compromisos:

Compromiso 1: asumo el doble compromiso metodológico-cognitivo computacionalista y enactivo, es decir, asumo que los seres humanos contamos con estos dos modos de cognición, lo que permitiría utilizar en ocasiones los métodos computacionalistas y en otras ocasiones los métodos enactivos. Considero que aceptar esto presenta ventajas frente a las explicaciones alternativas que sostienen un carácter excluyente e independiente.

Compromiso 2: asumo la arquitectura cognitiva PI, cuya estructura es capaz de soportar los dos modos de cognición, aunque ellos no compartan los mismos sustratos cognitivos. Desde esta perspectiva hay operaciones y estrategias socio-morales de distintos tipos. Por un lado, están las que surgen de los acoplamientos (sensoriomotores) dinámicos favorecidos por bucles de percepción-acción. Por otro lado, hay operaciones sujetas a procesamientos de información habilitados por representaciones y guiados por principios. Así, en esta propuesta sería posible en ocasiones *computar* y en otras *enactuar* la cognición socio-moral.

Es importante resaltar que el compromiso 2 (la arquitectura PI), posibilita el compromiso 1, lo que da pie a un tipo de pluralismo pragmático. De esta manera, ambos compromisos permiten un marco integrativo capaz de dar lugar a una propuesta que incluya tanto elementos computacionalistas como elementos enactivos, los cuales comprenden procesos distintos y ejecutan funciones distintas. A lo largo del capítulo presentaré con más detalle estos supuestos.

Ahora, doy paso a describir y explicitar la arquitectura cognitiva PI.

4.2. Arquitectura cognitiva PI

En esta sección revisaré los soportes materiales que están detrás de la arquitectura PI, pero antes de ello aludiré brevemente a dos temas: 1) a los elementos que habilitan mi propuesta, y, 2) a la ruptura que ha prevalecido entre las concepciones innatistas y de

aprendizaje en la ciencia cognitiva, lo que es el punto de partida de mi propuesta acerca de PI.

Respecto a (1), como se ha visto, mi propuesta deriva de conceder que hay dos tipos de modos cognitivos (computacionalista y enactivo), dos metodologías (individual y relacional) y dos entidades teóricas núcleo (representación y acción) que habilitan la cognición socio-moral humana. De ahí, entonces, he considerado que requerimos una arquitectura cognitiva capaz de soportar dos modos de cognición distintos⁷⁰.

Propongo que la arquitectura PI es capaz de cumplirlo, puesto que comprende tanto estructuras computacionales como estructuras *relacionales* de aprendizaje que se asemejan a los principios enactivos.

Ahora bien, el primer punto importante a mencionar es que en la historia de la ciencia cognitiva ha prevalecido una divergencia notable entre las concepciones innatistas de la cognición y el concepto de aprendizaje como tal. A continuación, considero este tema debido a que ha representado cismas teóricos importantes, pero a su vez, sus planteamientos son el punto de inicio de mi propuesta.

- **La divergencia entre el *aprendizaje* y lo innato tiene historia**

En 1948, Alan Turing planteó el tema del *aprendizaje* en los mecanismos computacionales, pues fue el primero en proponer la idea de que las computadoras pueden diseñarse simplemente por dejarlas “aprender” por sí mismas. En paralelo, Turing hizo explícito su compromiso empirista sobre la mente humana (Turing, 1948, p.16):

“Creemos entonces que hay grandes partes del cerebro, principalmente en la corteza, cuya función es en gran parte *indeterminada*. [. . .] Todo esto sugiere que la corteza del bebé es una máquina desorganizada, que puede organizarse mediante un *entrenamiento* de interferencia adecuado” (citado en Perconti y Plebe, 2020, p.6, las cursivas son mías).

Con ello, el autor se enfoca en reconocer la fuerte influencia que ejerce el *entrenamiento* en la mente humana, es decir, Turing otorgaba un papel sustancial al aprendizaje y al entorno para la conformación de la mente humana, pues al aludir a la idea de lo

⁷⁰ 1) Un modo cognitivo habilitado por procesamientos de información, representaciones, el uso de principios y que, además, soporte algoritmos programados para dar cuenta de lo innato, y, 2) un modo cognitivo habilitado por acoplamiento sensoriomotore agente-entorno favorecidos por bucles de percepción-acción, es decir, un modo cognitivo en donde la acción es guiada perceptualmente y surge del confrontamiento directo con el mundo.

indeterminado, claramente no los trataba como meros *desencadenantes* de lo que yace innatamente.

La idea de que la computadora “aprenda por sí sola” refiere a que el dispositivo puede contar con al menos dos tipos de arquitecturas: 1) una de base computacional que soporta procesamientos de información representacionales guiados por principios, así como también, algoritmos programados para efectuar ciertas tareas (rutinas), y; 2) una arquitectura capaz de soportar que, parte de su diseño, será aprendido en acción e interacción con el entorno mediante la *experiencia*, el reconocimiento de patrones, etc.⁷¹.

Desafortunadamente, este trabajo de Turing fue ignorado debido a que se le consideró un simple ensayo escolar y permaneció oculto durante décadas hasta que fue ratificado por Copeland y Proudfoot, en 1996 (Perconti y Plebe, 2020, p.6). Así que, mientras tanto, prevaleció una entera ruptura entre lo empirista y lo racionalista en el área computacional, en el conexionismo y en la inteligencia artificial en general, favoreciendo por completo el lado racionalista.

En paralelo, la ciencia cognitiva tradicional en sus orígenes sigue esta tendencia y se suscribe al lado racionalista. La tradición racionalista, que comenzó con René Descartes y fue seguida por Leibniz, Spinoza y Kant, entre otros, está en el centro de los lenguajes formales de la lógica y, en consecuencia, en la búsqueda de la construcción de modelos de razonamiento basados en símbolos y procesamientos de símbolos (Newell et al., 1957, 1959; Newell y Simon, 1972). De ahí, la metáfora computacionalista.

Así, pues, esto derivó en un marco de supuestos racionalistas adoptados en las metáforas cognitivas que implementan los modelos de la ciencia cognitiva, los cuales hasta hoy día comienzan a cuestionarse a partir de supuestos empiristas.

Para ver esto claramente, veamos a continuación los sustratos materiales detrás de PI.

4.2.1. ¿Qué es lo que habilita a la arquitectura cognitiva PI?

⁷¹ El área de Inteligencia Artificial ha evolucionado y en la actualidad no sólo se sabe que esto es posible, sino que esta idea se ha materializado y existen ejemplos que lo demuestran (p.ej. los procesadores que incluyen lenguajes de programación avanzada como *Machine learning* y *Deep learning*).

Como se ha visto, la arquitectura PI se compone de al menos dos arquitecturas, de lo que planteo que una se correspondería con el modelo computacionalista y la otra con el modelo enactivo.

La parte que se correspondería con el computacionalismo no presenta problemas, puesto que está basada en el modelo de procesamiento computacional tradicional que ha servido de soporte material para las posturas computacionalistas en la ciencia cognitiva, tal como se ha visto, en donde el cerebro se concibe como un “ordenador” y la cognición se entiende como un sistema procesador análogo a los sistemas artificiales y que, como ellos, “computa y ordena”, conforme a determinadas reglas, la información que recibe sobre el medio; proceso favorecido fundamentalmente por representaciones. En este contexto, el sistema cognitivo humano se concibe, esencialmente, como un dispositivo “manipulador de símbolos” y el pensamiento se concibe como la expresión dinámica de representaciones internas del mundo externo (Gutiérrez-Martínez, 2005).

Ello pone de relieve dos aspectos claves que resultan complementarios entre sí: el computacional y el representacional, de tal forma que la mente se concibe como estructuras de conocimientos, procesadores y representaciones, es decir, al igual que la máquina, la mente es concebida como un sistema interactivo de “conocimientos” y de “procesos” que manipulan “símbolos”, afirmación de la que deriva su modelo cognitivo: la entrada de información (p. ej., los *inputs* sensoriales) es procesada en función del conocimiento (almacenado) y mediante una serie de “reglas” preprogramadas (sistemas de producción) produce la actualización de una serie de modelos y representaciones, lo que efectúa bajo el control de un procesador o de varios procesadores (ver Gutiérrez-Martínez, 2005, p.21).

Esta metáfora informática de la mente se ha colocado en el centro de las discusiones entre los científicos cognitivos, sosteniendo que las mentes se entienden mejor como dispositivos computacionales y la cognición como procesamiento de información interna del sistema.

De esta manera, el computacionalismo busca “la redescrición formal de cualquiera y todos los sistemas en funcionamiento que se encargan del flujo de información bien formada” y, en este sentido, “lo que defiende es que cualquiera y todos los sistemas de información tienen que estar gobernados por «reglas» o procedimientos especificables que gobiernan lo que se hace con los inputs” (ver Gutiérrez-Martínez, 2005, pp. 155-156).

En particular, tal como se vio en el capítulo 1, la motivación de la metáfora computacional es que el computador clásico toma símbolos estructurados "similares al lenguaje" como entrada (*input*) y produce símbolos estructurados "similares al lenguaje" como salida (*output*). Este perfil entrada-salida (*input-output*) es mediado por la aplicación de reglas que en conjunto constituyen el programa que la máquina ejecuta.

Ahora bien, una acotación que debe ser hecha es que el modelo computacionalista tradicional asume que su funcionamiento es lineal, de abajo a arriba y ejecutado hacia adelante, es decir, enfatiza un marco de procesamiento jerárquico pero en serie, en donde la información siempre fluye en una sola dirección: la información sensorial viaja hacia arriba en la jerarquía, la cognición ocurre en ese procesamiento de información y luego la información motora viaja hacia abajo en la jerarquía en forma de comandos musculares (Carpendale, Frayn, y Kucharczyk, en Kiverstein, 2017, p. 198).

No obstante, como se hizo ver en el capítulo 2, una consecuencia de asumir la arquitectura súper-modular es que el procesamiento cognitivo-funcionalista ya no es lineal, de abajo a arriba y ejecutado hacia-adelante, sino que ahora es abierto, interactivo, multidireccional y se beneficia de múltiples recursos, lo que es enteramente compatible con la base computacionalista de la arquitectura PI.

Es compatible debido a que los avances de la IA han conducido a que los soportes materiales de los *programas informáticos*, si bien, se basan en la estructura computacional básica y tradicional, también en ellos se apela a estructuras más abiertas que involucran algoritmos no estáticos y por lo tanto, emplean representaciones igualmente no estáticas (o bien, representaciones superficiales, condescendientes e incompletas como las representaciones *SB* mencionadas en la nota al pie 41) y que se van conformando según las exigencias del entorno, además de que los procesos ya no son lineales, sino interactivos y multidireccionales, por lo que esta arquitectura es enteramente compatible con los planteamientos del enfoque súper-modular⁷².

⁷² Cabe mencionar que en realidad pienso que hay compatibilidades importantes entre la arquitectura súper-modular y los principios habilitadores de los *lenguajes de programación avanzada* de la IA (los cuales se verán más adelante), lo que quizás haría una compatibilidad más sólida entre el computacionalismo y el enactivismo. Sin embargo, para mostrarlo, habría que lograr al menos dos movimientos: 1) hacer compatibles las nociones de "representación" y "acción", y; 2) efectuar una fusión entre el racionalismo y el empirismo. Respecto a (1), considero que hay una oportunidad, debido a que el *procesamiento distribuido en paralelo* o paradigma PDP (el cual es una de las estructuras habilitadoras de los *lenguajes de programación avanzada* tales como *deep learning*), comparte la idea de representación mental, pero su formato es diferente a la representación

Así, pues, pienso que es plausible considerar que en esta estructura se basaría la parte computacional representacional de mi propuesta⁷³.

Por otro lado, la parte que procesa algo semejante a lo que postula el enactivismo proviene de la corriente empirista de la IA, para quienes, siguiendo a Hume, los conceptos son los productos de la experiencia y la razón obtiene todos sus materiales de la percepción. Esta parte es habilitada por el “lenguaje de programación avanzada” llamado *Deep learning* (DL en adelante), el cual tiene lugar principalmente por la invención exitosa de la *retropropagación*.

Previo a explicar esto, se debe destacar que en el campo de la inteligencia artificial (IA), no sólo encontramos explicaciones sólidas al respecto, sino que actualmente existen ejemplos

simbólica clásica, pues en PDP las representaciones se forman gradualmente, nunca son estables y sus modificaciones dependen de la experiencia, lo que encuentro compatible con la súper-modularidad. Respecto a (2), Joe Pater (2019) propone una fusión entre el empirismo y el racionalismo, pero aplicado al caso del lenguaje natural, en donde ofrece argumentos para la convergencia de la teoría chomskyana y el conexionismo, en particular, con las redes neuronales emergentistas que aceptan representaciones aprendidas. Considero que el desarrollo adecuado de (1) y (2), además de otros escrutinios más profundos, representa una oportunidad para lograr dicha compatibilidad, aunque en su elaboración habría que ser cautelosos para evitar colapsos conceptuales; trabajo que sale del alcance de la presente tesis, por lo que sólo lo menciono y su desarrollo quedaría para posibles investigaciones ulteriores.

⁷³ Aquí, se debe señalar que en la actualidad la base computacional de los PI se ha asociado con los principios conexionistas, sin embargo, dado que sus fundamentos relevantes (en origen) son compartidos con los del computacionalismo clásico, esto no representa problema, al menos para mis propósitos. La razón es que el conexionismo sigue concibiendo la cognición como procesamientos de información que utilizan representaciones para su ejecución. No obstante, debe mencionarse que estas representaciones son consideradas sub-simbólicas, y que el conexionismo ya no contempla el uso de reglas como tal para guiar la cognición. De ahí, entonces, son dos puntos que deben aclararse para evitar contradicciones: 1) el carácter sub-simbólico de las representaciones y 2) el no-uso de reglas. Concerniente a (1), las representaciones, bajo una lectura fuerte, son suplentes simbólicos que se mapean uno a uno en objetos externos y, en cambio, las representaciones sub-simbólicas son estructuras (sub-simbólicas) mapeadas en elementos externos, es decir, las estructuras sub-simbólicas se conforman de múltiples componentes, los cuales pueden figurar o no en otras múltiples representaciones, mismas que a su vez pueden incluir componentes adicionales (De Bruin y Kastner, 2012, p. 552, nota al pie 12). El punto a destacar es que ambas concepciones implican la portación de información, y debido a que la condición mínima para que un objeto mental cuente como representación es que ésta sea una estructura portadora de información, por tanto, no hay incompatibilidad. En segundo lugar, como señala Joe Pater, “no hay nada en un modelo conexionista que prohíba el uso de símbolos, incluidas las variables, y otras representaciones desarrolladas en las tradiciones lingüísticas” (Pater, 2019, p.54). Así, pues, se trata de ampliar el alcance de las representaciones y ya no pensarlas rígidamente estáticas, sino que se van conformando según el contexto; lo que es compatible con la súper-modularidad. Y, concerniente a (2), algunos autores han sostenido que es falso que el conexionismo no permita el uso de reglas, pues las funciones que constituyen sus algoritmos son reglas, incluso las dinámicas de las funciones sub-simbólicas pueden traducirse e intercambiarse con las reglas simbólicas de los sistemas del computacionalismo tradicional (Gutiérrez-Martínez 2005, p. 294). En consecuencia, punto (1) y punto (2) en realidad no presentarían inconsistencias.

materiales de este tipo de arquitecturas, los cuales, por antonomasia, son los lenguajes de programación avanzada llamados *Machine learning* y *Deep learning*. Ellos habilitan sistemas que aceptan que gran parte del diseño de la máquina sea mediante el *aprendizaje* (incluso desde cero) proveniente de las acciones e interacciones con el entorno, mismo que puede llevarse a cabo mediante la experiencia. Para mis propósitos, me enfoco sólo en *Deep learning* (DL).

En particular, DL se basa en las redes neuronales artificiales y en el *procesamiento distribuido en paralelo* (PDP), que es significativamente capaz de funcionar en modo de *retropropagación*, en donde el aprendizaje y la acción son sus fundamentos. En breve, se trata de lo siguiente:

“La estructura básica de la maquinaria “distribuida en paralelo” está formada por unidades simples organizadas en distintas capas, con conexiones unidireccionales entre cada capa y la siguiente. Esta estructura, conocida como red *feedforward*, se conserva en la mayoría de los modelos de *deep learning*. PDP se adhiere a una cuenta empirista radical, con modelos que aprenden cualquier función significativa posible desde cero, solo por la experiencia. El éxito de PDP se debió en gran medida a una regla matemática eficiente, conocida como *retropropagación*, para adaptar las conexiones entre unidades, a partir de ejemplos de la función deseada entre la entrada y la salida conocidas”. (Perconti y Plebe, 2020, p. 7)

Concretamente, se trata de hablar de redes neuronales artificiales que avanzaron gracias a los desarrollos de PDP, el cual tiene como estructura una red de retroalimentación (*feedforward*) y se basa en la regla matemática de *retropropagación* (*i.e.*, que actúa hacia atrás) lo que permite el *aprendizaje* y en ello se encuentra su éxito. Esta estructura es preservada en DL y funge como su principal habilitador.

Los detalles de sus sustratos materiales, sus funcionalidades y enlaces, sus capas habilitadoras, entre otros temas, son materia de debate en la IA, lo que sale del alcance de la presente tesis (para una discusión minuciosa respecto a DL y la ciencia cognitiva corporizada ver Perconti y Plebe, 2020; y, para un análisis profundo de la estructura que constituye a los lenguajes de programación avanzada ver Raschka y Mirjalili, 2017).

Para mis propósitos, es suficiente con notar que el *aprendizaje* de DL puede ser desde cero y que surge a partir de la experiencia obtenida de la *acción* e *interacción* con el entorno, con lo que presenta ventajas ecológicas que se ajustan bien a los principios del enactivismo.

Así, pues, si mis consideraciones son correctas, las estructuras vistas constituirían la arquitectura cognitiva PI, las cuales posibilitan proponer la integración de los modelos cognitivos computacionalista y enactivo, respectivamente.

En torno a ello, considero relevante destacar que en la actualidad encontramos explicaciones sólidas y ejemplos materializados del tipo de arquitectura aludido PI, los cuales habilitan sistemas extraordinarios que intentan asemejar la cognición humana en diversas tareas, mismos que se favorecen de las dos estructuras vistas, *i.e.* involucran la base computacional, pero también, *lenguajes de programación avanzada* tal como DL (y en mi interpretación, los principios computacionalistas y enactivos).

Un ejemplo de ello que vale la pena mencionar es el llamado *AlphaGo* (y sus versiones derivadas *AlphaZero*, *AlphaGoZero* y *MuZero*), que es un PI con base computacional y habilitado por DL, hecho para jugar el juego de mesa llamado *Go*. Este juego *Go* es de origen chino y se ha considerado uno de los más complejos de la historia, debido a que involucra niveles muy abstractos de estrategia y exige altos niveles de creatividad, entre otras múltiples aptitudes.

En particular, *AlphaGo* y sus distintas versiones aprenden mediante la (percepción-) acción y la experiencia que de ello surge; y su función puede habilitarse de diversas maneras: en ocasiones empleando reglas y representaciones y en otras no, a veces haciendo uso de procesamientos de información para ejecutar la acción y otras no, y en su lugar, responden al entorno sin mediaciones; características que considero compatibles con los dos modos de cognición que planteo.

Por ejemplo, *AlphaGo* cuenta con las reglas del juego y aprende a usarlas jugando miles de partidas, es decir, aprende mediante la experiencia que obtiene de jugar repetidamente y del uso de reglas que le guía; *AlphaZero* tiene sólo las reglas básicas del juego y a partir de su uso continuo las expande y aprende; *AlphaGoZero* se considera el más creativo, debido a que es capaz de *aprender desde cero*, pues descubre nuevos conocimientos, nuevas estrategias y nuevos movimientos a partir de jugar contra sí mismo y de juegos aleatorios, y *MuZero* sorprendentemente aprende sin reglas, *i.e.* no posee ningún tipo de regla, sino que a partir de la *acción* aprende un modelo de su entorno, lo que le permite planificar estrategias ganadoras en dominios desconocidos; esta versión se dirige a ser un sistema de aprendizaje de propósito general.

En adición, no sobra decir que en el año 2015 *AlphaGo* derrotó por primera vez al campeón mundial del momento del juego *Go*, convirtiéndose en la primera máquina en derrotar a un profesional humano de este juego. Este dato es relevante debido a que marca un hito en el campo de la inteligencia artificial, puesto que por primera vez existe un dispositivo que es capaz, no sólo de asemejar la creatividad humana, sino de superar su pensamiento estrategia y de descubrir en tiempo récord nuevos conocimientos y nuevos movimientos, lo que logra por ser capaz de acumular miles de años de conocimiento humano en tan solo unos pocos días.

Así, pues, si mis interpretaciones son correctas, el esbozo de estos ejemplos puede ilustrar el potencial que posee la arquitectura PI, lo cual, considero, puede brindar nuevas y posibles vías de exploración para el campo de la ciencia cognitiva.

Ahora bien, tras revisar los sustratos materiales de PI, a continuación, planteo esta arquitectura como metáfora cognitiva.

4.2.2. Arquitectura PI como metáfora cognitiva

La inteligencia artificial ha sido el “corazón intelectual” de la ciencia cognitiva, como le ha llamado Boden (2006). La razón es que las arquitecturas cognitivas procedentes de los postulados de la IA muestran un progreso irrefutable, lo cual ha nutrido significativamente el campo de las ciencias cognitivas al dotarlas con argumentos que implementan soportes materiales capaces de corresponderse con sus planteamientos relativos a los aspectos psicológicos de la cognición, contribuyendo así a las perspectivas computacionalistas, conexionistas, de redes neuronales artificiales y de sistemas dinámicos, entre otros.

Para mis propósitos, lo relevante es notar que, si concedemos la idea de que la cognición humana puede corresponderse con los planteamientos de las dos arquitecturas aludidas por PI, entonces, el *aprendizaje* (no programado o no-predeterminado) en realidad no estaría peleado con las arquitecturas cognitivas que implican nociones de módulos⁷⁴, de nodos o de conexiones que, incluso, se comprometan con la idea del desarrollo genéticamente determinado para asegurar la universalidad de trayectorias; tampoco estarían peleadas las nociones de representación y acción en los términos definidos; y

⁷⁴ Este planteamiento acerca del aprendizaje no predeterminado también es sostenido por la súper-modularidad.

tampoco se negaría el posible uso de las dos metodologías (interna y relacional) participando en una misma arquitectura cognitiva.

Me explico; aquí se trata de asumir el doble compromiso metodológico-cognitivo computacionalista y enactivo, ante lo cual he propuesto que la arquitectura cognitiva PI es capaz de soportarlo. En este contexto, pienso que debido a que hoy contamos con explicaciones robustas acerca de los soportes materiales que le habilitan y que ellos, al parecer, tienen el potencial de corresponderse con los aspectos psicológicos postulados por el computacionalismo y por el enactivismo, tal como se ha visto (al menos desde un mapeo de *grano grueso*), entonces tenemos buenas razones para pensar que una arquitectura del tipo PI es capaz de dotar con argumentos coherentes una posible postura integrativa que comprenda rasgos cognitivos tanto computacionalistas como enactivos, *i.e.* una postura de la cognición computacionalista-enactiva.

En concordancia con ello, entonces, es posible sugerir que una parte de la cognición socio-moral humana se correspondería con los procesamientos computacionalistas y otra parte con los procesamientos enactivos, respetando los principios que cada uno sostiene. De esta manera, la parte computacional sería capaz de soportar la arquitectura súper-modular y la parte enactiva soportaría la arquitectura que se corresponde con sus principios relacionales de acción e interacción entre el agente y el entorno.

Así, hoy considero razonable recurrir a los ejemplos aludidos, pues estos emulan de manera muy cercana a la cognición humana. Sin embargo, se debe mencionar que, en principio, los lenguajes de programación avanzada, tales como DL, se han creado sólo con objetivos ingenieriles y no con fines de explorar o investigar la cognición humana, lo que no es razón para socavar la gran contribución que podrían brindar a la ciencia cognitiva (ver Perconti y Plebe, 2020) y para mis propósitos, para sostener y robustecer mi propuesta.

Aquí el objetivo es escudriñar cómo funciona la cognición socio-moral humana, frente a lo cual, propongo que tenemos buenas razones para pensar que es mediante una parte súper-modular que aseguraría la universalidad de la moral humana a la vez que involucra información pragmática, pero por otra, mediante una parte (no representacional) capaz de aprender a partir de las meras acciones con el entorno, lo que haría justicia al aspecto relacional y situacional de la misma e incluiría la intervención de los procesos supra-personales autónomos más allá de las representaciones individuales.

El punto de partida sería conceder que la mente humana es capaz de hacer todo esto y mucho más, y si esto es así, entonces, ¿por qué restringir las posibles explicaciones? (Premack, 2004). En su lugar, propongo que sería fructífero ampliar los marcos disponibles y recurrir a posturas que permitan superar las dicotomías prevaecientes y con ello pensar en trabajos conjuntos entre los diversos marcos conceptuales.

Por último, no sobra aclarar que esta arquitectura PI no es la misma que aquella propuesta por los teóricos de la cognición social que se vio en el capítulo 2 (sección 2.5.2), con la que se busca la hibridación atribucionista-interaccionista mediante un sistema dual que apela a una lógica mecanicista à la Bechtel.

4.2.3. Distinción entre PI y la arquitectura dual socio-cognitiva

En breve, la arquitectura que proponen estos autores (Bohl y van den Bos, 2012; Gangopadhyay y Schilbach, 2011) se basa en las teorías de sistemas duales de la cognición social, que permiten hablar de dos sistemas socio-cognitivos: el sistema 1 encargado de los procesamientos automáticos, rápidos y sin esfuerzo y, por otra parte, el sistema 2 encargado de procesamientos deliberativos, voluntarios, más lentos y que implican esfuerzo. En adición, esta arquitectura apela a la naturaleza mecanicista del tipo Bechtel (2008), que es capaz de soportar múltiples capas para diferentes niveles de explicación. Así, los autores postulan, a partir de un mapeo de *grano-grueso*, que el sistema 1 se correspondería con la metodología interaccionista y el sistema 2 con la metodología atribucionista, y esto, asociado con la lógica mecanicista del tipo Bechtel, permite diferentes capas explicativas que dan cuenta de los diversos niveles cognitivos, tales como el sub-personal, el personal y el supra-personal.

Ahora bien, si ya contamos con este modelo cognitivo dual de la cognición social, cabe preguntar ¿por qué recurrir a la arquitectura PI?

4.2.4. Justificación de la pertinencia de PI

Mi respuesta a modo de justificación es la siguiente: primero, estoy completamente de acuerdo con esta arquitectura dual que busca establecer un marco integrativo de la cognición social, puesto que parto de ella para establecer la necesidad de ambas metodologías, no obstante, cabe mencionar que esta arquitectura se basa en las explicaciones tradicionales de ToM, las cuales asumen una arquitectura computacionalista clásica dedicada (en su mayoría) a procesamientos de pensamiento abstracto y

deliberativo, *i.e.*, lo que se conoce como cognición de alto-nivel. En contraste, en mi propuesta recorro a la arquitectura súper-modular, movimiento que derivaría en un colapso con el sistema 1.

Me explico, en primer lugar, mi tema no sólo es la cognición social sino la cognición socio-moral, lo que amplía el marco de debate. En torno a ello, mi propuesta es que la cognición socio-moral es habilitada por la arquitectura computacionalista súper-modular y por la arquitectura enactiva. El problema surge en lo concerniente a la parte computacionalista, debido a que los principios de la súper-modularidad colapsarían con el modelo, pues si bien la súper-modularidad es de naturaleza computacionalista, ésta no se correspondería exclusivamente con aquellos procesamientos del sistema 2, sino que también se correspondería con aquellos del sistema 1; la razón es que la súper-modularidad postula fundamentalmente la competencia moral como un *saber-cómo* (con su correspondiente *saber-qué*), tal como se vio. En particular, recordemos que el saber-cómo súper-modular apela al uso de heurísticas rápidas y frugales, el cual es ejecutado en tiempos de reacción medibles en milisegundos, lo que da pie a la competencia moral automática y sin esfuerzo, lo que podría corresponderse con lo postulado en el sistema 1. Bajo esta lectura, el colapso radicaría en que la concepción de la estructura dual otorga un papel *limitado* al computacionalismo, pues éste habilita exclusivamente el sistema 2 y el interaccionismo el sistema 1, mientras que, en mi propuesta, el computacionalismo tiene pertinencia tanto en el saber-qué moral como en el saber-cómo moral, ambos habilitando (en interacción) la competencia moral.

Por ello, si adoptara dicha arquitectura dual de la cognición social, ocurriría un colapso conceptual que limitaría el alcance de la explicación del tema de interés y con ello dificultaría distinguir la importancia de cada marco teórico y el papel que juega cada uno de ellos en mi propuesta integrativa, razón por la cual se justifica la pertinencia de la arquitectura PI, o al menos así lo defenderé.

En adición, aquí no se está hablando de un sistema dual sino de dos modos cognitivos: uno guiado por principios y el otro guiado perceptualmente, cada uno con sus sustratos cognitivos habilitadores. En este punto, se debe mencionar que, si mis consideraciones son correctas, lo que legitima la pertinencia de ambos modos cognitivos es aceptar que los dos tienen lugar en la competencia moral: el que recurre al uso de razonamiento heurístico guiado por principios y favorecido por representaciones (no necesariamente acabadas), y el ultrarrápido guiado perceptualmente, el cual no apela al uso de representaciones y/o

reglas, sino que emerge de la interacción directa con el entorno. Esto lo sugieren los estudios que muestran los tiempos de reacción (en milisegundos) que cada uno sostiene. Por supuesto, delimitar exactamente cuáles tareas morales son habilitadas por uno y cuáles por el otro, requerirá una investigación empírica mucho más profunda, sin embargo, se debe hacer notar que ambos modos cognitivos se enfocan en distintos aspectos de la cognición socio-moral, según sus fundamentos: el enactivo, a partir de su concepción de agencia e interacción, se enfoca en los aspectos fenomenológicos, interactivos y corporizados, y el computacionalismo, por su parte, en los procesamientos de información (cargados de contenido y guiados por principios) que en su interacción dan lugar a la cognición moral.

Ahora bien, una vez esbozada la arquitectura cognitiva PI que propongo y de haber trazado su posible distinción con aquella arquitectura dual de la cognición social, ahora corresponde abordar los cuatro temas en los que habría incidencia entre el computacionalismo y el enactivismo, según mi interpretación: 1) evolución, 2) cognición, 3) cognición social y 4) cognición moral, y de ahí, veremos sus implicaciones normativas.

4.3. Evolución

En las secciones anteriores dedicadas al argumento evolutivo (secciones 2.1 y 3.1), abordé los supuestos que sostienen la escuela de Santa Bárbara y la hipótesis de la construcción de nicho, respectivamente. En ello, se vio que la escuela de Santa Bárbara resalta el modularismo, el adaptacionismo y reconoce la selección natural como primera fuerza evolutiva, en donde se defiende que los organismos pasivamente se adaptan (o no) a su medio y en este proceso los organismos son afectados por el entorno.

En cambio, la hipótesis de la construcción de nicho, si bien está anclada a la evolución biológica, también sostiene que los organismos interactúan activamente con su entorno, en donde el entorno y los organismos se transforman y moldean mutuamente, es decir, se co-constituyen, por lo que desde este punto de vista se otorga un papel *activo* a los agentes. Aquí, se vio una aproximación a una reformulación del término de *adaptación*, en donde se reconoce su carácter dual y se plantea la *acomodación* del organismo al entorno, pero también la *asimilación* del entorno por el organismo (Álvarez, 2013).

Así, se planteó que la selección natural pudo propiciar la evolución de un núcleo biológico de rasgos heredables que incluiría instintos sociales, paternos, filiales y empatía, integrados

con ciertos mecanismos que subyacen a nuestras orientaciones normativas generales, todo ello impulsado principalmente por la vulnerabilidad corporal del hombre.

Y, concerniente al nicho cultural humano, que es habilitado por el *aprendizaje social*, se reconoció la importancia de la aloparentalidad, de la intersubjetividad, de la comunicación simbólica y de las condiciones corporizadas que atraviesa a cada una de ellas, pues estos serían los factores (culturales) que contribuyeron a configurar el marco para la emergencia de las conductas cooperativas y prosociales en el hombre.

De ello, entonces, el primer desafío es articular coherentemente los principios de la escuela de Santa Bárbara (herencia genética) con los principios de la construcción de nicho (herencia ecológica); al menos una aproximación de *grano grueso* que permita habilitar el argumento del conocimiento moral general como dotación evolutiva, para lo cual recorro principalmente a los argumentos de la genómica cultural.

La genómica cultural se basa en la hipótesis de la *herencia dual* postulada por Richerson y Boyd, la cual se enfoca principalmente en la evolución cultural, pero se diferencia de ésta debido a que la genómica cultural busca explicar, no sólo la evolución cultural, sino la genuina interacción entre los genes y la cultura, es decir, se enfoca en la *influencia recíproca* prevaleciente entre ellos para mostrar cómo los genes influyen en la cultura y cómo la cultura influye en los genes, lo que intenta corroborar empíricamente.

De esta manera, la genómica cultural trata de superar el modelo estático de la evolución biológica que había sido la base de las discusiones tradicionales y ahora, en su lugar, postula que el genoma humano en realidad es *maleable*, lo que abre un abanico de posibilidades pertinentes a los relatos evolucionistas.

De ello, en este apartado me enfoco en los siguientes temas: en primer lugar, intento aproximarme a una explicación plausible que aborde la convergencia co-constituyente entre lo biológico y lo cultural mediante los principios de la genómica cultural. En segundo lugar, se verá que hay una gran aportación a ello si se considera la pertinencia del concepto del *cuerpo vívido*, pues éste se postula como el punto de inflexión entre ellos. En tercer lugar, enfatizaré que un rasgo primordial del aprendizaje social que se logra por *imitación*, es que en éste se reconoce al otro como agente intencional, pero también, que la acción que se imita posee significado en sí misma, pues en la imitación se atribuye un significado o sentido objetivo a la acción, pero que a su vez permite múltiples interpretaciones. Esto se muestra

cuando el agente imita una acción que, aunque no comprenda su significado o intención, se llega al resultado (exitoso) que se espera; lo que posiblemente pudo ocurrir con las estrategias exitosas y perniciosas de nuestros antepasados. Así, pues, si mis consideraciones son correctas, estos planteamientos en conjunto permiten argumentar a favor de aquellas orientaciones normativas que constituyen un tipo de conocimiento moral dotado evolutivamente, con su carácter general y situado.

4.3.1. Convergencia bio-cultural

Frente a la dialéctica prevaleciente entre la evolución biológica y la evolución cultural, considero que es posible legitimar las ideas planteadas si, en primer lugar, se reconoce que:

“Para cualquier escala de tiempo dada y cualquier nivel de análisis, existen procesos verdaderamente interactivos entre la cultura y los genes. En todo caso, los procesos culturales son quizá la fuerza impulsora más fuerte detrás de la evolución humana reciente (Chen et al., 2016). Con esa perspectiva corregida, la biología ya no es de temer sino algo con lo que trabajar. Necesitamos comprender hasta qué punto los contextos eco-culturales locales han dado forma a los genomas de diferentes grupos y cómo interactúan con los cambios culturales” (Chen y Moyzis, 2018, p. 780).

En torno a ello, se han realizado esfuerzos para explicar las interacciones entre los genes y la cultura, en donde se reconocen algunos ejemplos paradigmáticos que podrían enlistarse de la siguiente manera, según su unidad de análisis:

- 1) A nivel de “toda la especie”, los ejemplos paradigmáticos son el lenguaje y la auto-domesticación. El *lenguaje* parece ser el caso más célebre de la influencia mutua entre los genes y la cultura, debido a que se reconoce una fuerte presión de selección asociada a las mutaciones recientes del gen FOXP2 (el gen asociado al lenguaje), presión mantenida por la dependencia de la cultura humana en el lenguaje para la vida en grupo⁷⁵.

⁷⁵ En concreto: “En el caso de la evolución del lenguaje, el gen FOXP2 es el caso más célebre. Su papel en el desarrollo del habla y el lenguaje normal (particularmente la gramática) ha sido bien documentado. Dos mutaciones en FOXP2, en las posiciones 911 y 977 en el exón 7, separaron tanto a los humanos como a los neandertales de otros primates (Krause et al., 2007). Se cree que estas mutaciones prepararon el escenario para el surgimiento del lenguaje humano, lo que sugiere que las mutaciones genéticas condujeron al elemento cultural posiblemente más importante de todos: el lenguaje. Por supuesto, es probable que FOXP2 sea una condición necesaria, pero de ninguna manera suficiente para el desarrollo del lenguaje. Varios otros genes se han implicado en el procesamiento del lenguaje [...]. Se ha encontrado que algunos de ellos muestran signos de selección reciente (p. ej., FOXP2, ROBO1, ROBO2, CNTNAP2 y KIAA0319; Mozzi et al., 2016), quizá debido a una fuerte presión de selección presentada por la dependencia de la cultura humana en el lenguaje para la

La *auto-domesticación*, por su parte, se muestra por las grandes tendencias que ha tenido el humano a formar grupos cooperativos más grandes, a ambientes sedentarios, a la conformidad, a la moralidad y a la co-residencia con individuos no familiares (ver Chen y Moyzis, 2018, p. 769); lo que ha resultado en una cascada de progresos evolutivos físicos y psicológicos en el hombre.

- 2) A nivel “específico de un grupo”, es un desafío determinar si la adaptación genética específica a un grupo está asociada con las diferencias culturales, no obstante, se han propuesto como ejemplos paradigmáticos de ello: “la migración y sus implicaciones”, “la relación del transportador de serotonina con las orientaciones culturales colectivistas e individualistas” y por último “la inscripción universitaria y los neurotransmisores” (para una revisión detallada de ellos ver Chen y Moyzis, 2018, pp. 771-773).

Con estos ejemplos, se ha buscado reconocer que la diversidad biológica es tan necesaria como la diversidad cultural, y ambas deben valorarse para explicar cualquier fenómeno evolutivo humano, puesto que “si uno se enfoca en una sola de las dos variaciones [...], pero no en ambas, es probable que cualquier conclusión sobre las diferencias de grupo sea incompleta e incluso engañosa” (Murdoch et al., 2013).

El planteamiento incluye casos controversiales que derivan en ideas de eugenismos, racismos, creencias de razas superiores e inferiores, etc., casos en los que se toma en cuenta una sola variación, ya sea la genética o la cultural, y conducen a nociones discriminatorias, marginales y de segregación.

Por estas razones se precisa reconocer la influencia recíproca entre genes y entorno (cultural), y de esta manera superar la noción de una causalidad unidireccional en que los ambientes filtran los genes de ascendientes a descendientes y, en su lugar, reconocer un proceso de presión evolutiva co-constituyente entre el entorno y el organismo, en el que habría afectaciones y compensaciones entre los genes y la cultura. Por supuesto, la temporalidad de evolución entre lo biológico y lo cultural es dimensionalmente distinta, no obstante, en este contexto se reconocerían las constantes influencias entre ellos, aunque indudablemente con escalas temporales diferentes.

vida en grupo. El lenguaje humano parece ser el resultado de la influencia mutua entre los genes y la cultura”. (Chen y Moyzis, 2018, p. 769).

Asimismo, se busca superar el punto de vista tradicional en el que se piensa que las interacciones gen-cultura se efectúan en una sola dirección: de los genes a la conducta y por tanto a la cultura (genes → conducta → cultura), lo que ha resultado en determinismos biológicos que sólo reconocen la influencia de los genes en la cultura, con lo que pasan por alto la diversidad cultural y sus influencias en la biología. Habilitar estas ideas es posible si se toma en serio el objetivo general de la genómica cultural:

“En última instancia, la genómica cultural tiene como objetivo integrar la historia de los genes (evolución molecular, incluida la introgresión), las funciones bioquímicas, los mecanismos epigenéticos, los correlatos de comportamiento (especialmente los fenotipos refinados) y la presión de selección que plantean los elementos culturales, así como sus mecanismos (naturales, selección sexual o social).” (Chen y Moyzis, 2018, p.780).

De esta manera, sería posible pensar que las condiciones otorgadas tanto por la selección natural como por la construcción de nicho, en su continua interacción e influencia recíproca, constituyeron el marco definicional que dio lugar a la emergencia de un tipo de conocimiento moral general conformado por aquellas orientaciones normativas que surgieron de las estrategias exitosas o perniciosas de nuestros antecesores, las cuales componen nuestras nociones amplias y generales pertinentes a lo que consideramos bueno, malo, correcto, incorrecto, etc.

El planteamiento se articularía de la siguiente forma: el núcleo biológico de rasgos heredables dotado por la selección natural, no sólo originaría el nicho cultural que propició la emergencia de las capacidades prosociales humanas, sino que habría genuinas influencias recíprocas entre estos dos tipos de herencia (genética y ecológica), de tal suerte que mediante la interacción de dos rutas lógicamente diferentes se posibilita un ajuste capaz de generar un nicho dinámico bio-cultural para la emergencia de tal conocimiento moral general.

Se trata, pues, de pensar en un relato evolucionista de influencia recíproca (así definida) entre estos elementos, pues como se hizo ver, las prácticas de aloparentalidad catapultaron y llevaron la cooperación humana a otro nivel, pero a su vez, la aloparentalidad fue posible gracias a la constitución humana ya dotada de empatía y de aquellos instintos filiales, parentales y de socialidad, mismos que propiciaron, en primer lugar, el marco en el que se gestó la vida en grupo. Estas características entrelazadas con influencias recíprocas entre sí compondrían el nicho dinámico biológico y cultural que favoreció el progreso de las conductas cooperativas, pro-sociales, de ayuda, y en última instancia, de nuestras

orientaciones normativas generales, en donde las intervenciones de la imitación, de la comunicación simbólica y de la intersubjetividad resultaron fundamentales.

Así, nos aproximamos a una descripción de *grano grueso* pertinente a los factores que pudieron beneficiar la emergencia de aquel tipo de conocimiento moral que planteo, el cual formaría parte natural de nuestro aparato cognitivo.

No obstante, aún falta ver ¿qué contribuyó exactamente a generar aquellas nociones amplias y generales acerca de lo “correcto” o “incorrecto” que constituyen tal conocimiento moral?

A continuación, abordo directamente las posibles condiciones que pudieron favorecer la emergencia de esas nociones normativas generales, las cuales son amplias en naturaleza y beneficiadas por condiciones corporizadas, en donde la pertinencia de la *imitación* y del concepto de *cuerpo vívido* entra en juego.

4.3.2. El *cuerpo vívido* como punto de inflexión entre lo biológico y lo cultural

Gregor Etzelmüller sugiere que el concepto de *cuerpo vívido* (*Leib-Körper*) proporciona un puente genuino entre la naturaleza y la cultura, entre la antropología evolutiva y la antropología histórica. En particular, la distinción clásica entre cuerpo (*Körper*) y cuerpo vívido (*Leib*) se remonta a Edmund Husserl (1859-1938), quien sostiene que, si bien, podemos convertir el cuerpo en un objeto, el *cuerpo vívido* como tal evade esta objetivación. La razón es que el *cuerpo vívido*, en su aspecto objetivo y subjetivo, es naturaleza, pero también es un ser capaz de trascender la naturaleza en cultura. Con estas ideas, Etzelmüller busca denotar la ambigüedad del concepto del cuerpo vívido y establecerlo como “el punto de inflexión” entre la naturaleza y la cultura⁷⁶.

Como punto de partida, Etzelmüller considera el hecho de que el desarrollo de los seres humanos es lento y prolongado, por lo que depende en gran medida de los factores ambientales. Esto hace surgir al humano como una entidad caracterizada por una interdependencia mutua e irreductible entre naturaleza y cultura, cuerpo y psique. La razón es que los hechos biológicos están involucrados de manera crucial en la conformación

⁷⁶ Nótese que esta idea es compatible con los planteamientos enactivos concernientes a cómo los agentes humanos nos auto-individuamos del mundo; movimiento con lo que logramos una perspectiva des-centrada que da lugar, entre otras cosas, a una perspectiva moral del mundo.

social y psíquica del hombre en tanto que la “baja velocidad” de su desarrollo somático da lugar a una potencialidad biológica que es considerada como un tipo de “plasticidad abierta” moldeada fundamentalmente por el entorno y, debido a esta plasticidad del desarrollo humano, existiría una perpetua retroalimentación entre la dinámica cultural y la biología humana, de tal manera que “los humanos somos seres culturales por naturaleza” (Etzelmüller, 2016), lo que es compatible con los principios de la genómica cultural.

Lo que motiva el planteamiento es pensar la cultura como intrínseca al ser humano, en tanto que son los procesos culturales y el aprendizaje cultural las primeras herramientas que dan soporte al desarrollo biológico y social de los infantes humanos. Así, la potencialidad biológica, considerada como *plasticidad abierta*, estaría enraizada en nuestra dotación evolutiva y a su vez atravesada y moldeada genuinamente por el entorno socio-cultural.

Por lo que, según Etzelmüller, no existiría la mera biología de la vida humana, pues la naturaleza en origen adquiere una dinámica cultural. En esta línea, se puede pensar coherentemente en la propensión natural humana a la moralidad; siendo ésta un fenómeno cultural pero enraizado en la historia natural.

Pero, ¿cómo se enlazan concretamente estas ideas con la imitación y con la emergencia del conocimiento moral general?

Como se vio en la sección de construcción de nicho (sección 3.1), la evolución de la intersubjetividad humana fue favorecida primordialmente por las prácticas de aloparentalidad y por la imitación corporal; factores determinantes que promovieron la cooperación y la pro-socialidad entre individuos. En este contexto, se arguye que la adaptación social y cooperativa humana se benefició fundamentalmente de la *imitación excesiva*, la cual denota grados extremos de intencionalidad compartida (Tomasello, 2010, p.154; Searle, 1997).

La razón es que Etzelmüller sostiene que la *imitación excesiva* fue un rasgo clave para que el *aprendizaje cultural* tuviera lugar (mismo que acelera el desarrollo cultural), pues se ha considerado que “la imitación está involucrada en la *atención compartida*, lo que revela un modo de ser fundamentalmente social en el mundo y sienta las bases para la comunicación simbólica” (Etzelmüller, 2016, p. 198). Estos aspectos se han considerado decisivos, no sólo para habilitar el aprendizaje cultural, sino también para la evolución del lenguaje, de la moral y del pensamiento humano en general. De esta manera, el autor arguye que la

imitación (que es esencialmente corporizada) es el *punto de inflexión* en el que la evolución natural se traslada al desarrollo cultural.

Estrictamente hablando, Etzelmüller sugiere que la imitación sería el puente que une la razón, el sentido y el significado de la intencionalidad con las realidades del mundo, pues mediante la imitación los humanos comparten e intercambian significados; por ejemplo, considérense los rituales, pues al participar en los rituales descubrimos y creamos formas que permiten compartir significados entre humanos.

El argumento se basa en la hipótesis de Bernhard Waldenfels (2016), quien defiende que la imitación tiene en sí misma un sentido o significado intrínseco. Para sostenerlo, Waldenfels entrelaza inextricablemente los conceptos de “*pathos*” y de “*respuesta*”, en donde “*pathos*” denota dos aspectos: 1) la sensibilidad que tenemos por lo que hace el *otro*, y, 2) su interpretación; pero esta interpretación no es un simple proceso causal ni una creación libre sin “*fricción*” con la realidad, sino que se corresponde con ella. Y, por su parte, la “*respuesta*” sería la articulación vocal o física que efectuamos en correspondencia con lo que hace el *otro*; aquí debe resaltarse que esta respuesta toma su orientación del exterior, por lo que su éxito depende de factores externos y contextuales.

De ello, el planteamiento es que prevalece un entrelazamiento corporizado entre *pathos* y *respuesta* en la imitación, el cual crea un puente entre la razón y las realidades objetivas (aunque multivalentes) del mundo cuando el sujeto efectúa la imitación en correspondencia con lo que percibe del otro, incluso cuando la imitación se efectúa como un “acto de fe” que supone significado a la acción que se imita.

Esto significa que la imitación es capaz de capturar un aspecto *objetivo* pero *multivalente* de la acción (y de la intencionalidad). Bajo esta lectura, el aspecto *objetivo* concierne al significado tocante a la realidad, y el aspecto *multivalente* concierne a la multiplicidad de interpretaciones que le subyacen.

Por ejemplo, la imitación de prácticas o rituales supone significado en sí mismo, lo que incluso podría llamarse un “acto de fe”, puesto que, aunque se les adscriba un sentido objetivo, su imitación permite una amplia variación de interpretaciones, es decir, las acciones que se imitan pueden entenderse de maneras muy distintas, pero siempre se reconoce que su intencionalidad tiene un sentido o un significado objetivo. Considérese, por ejemplo, la cooperación humana en una cacería, en la cual los participantes pueden

entender el proceso en términos muy diferentes, tales como "ganar el pan" o como "un freno al caos de la naturaleza" o como "camaradería", siempre que el esfuerzo y el resultado sean lo suficientemente plausibles y exitosos (Etzelmüller, 2016). En otras palabras, el sentido y el significado objetivo de la intencionalidad y de la acción se mantienen en la imitación, pues lo que se imita tiene un porqué, un sentido o una razón objetiva y multivalente.

Para sumar más apoyo a estos planteamientos, Wulf (2016) argumenta que el conocimiento práctico y la actividad social encuentran su punto de encuentro también en la *imitación*, puesto que ambos están moldeados en gran medida por factores históricos y culturales que se transmiten fundamentalmente mediante la *mímesis*⁷⁷, lo que a su vez habilita el aprendizaje cultural y reconoce la intencionalidad con un sentido y significado objetivo.

En este punto, cabe agregar que hay trabajo documentado acerca de cómo los infantes humanos imitan tanto las señales relevantes como las señales irrelevantes de un acto (Etzelmüller, pp. 209-210). Esto sugiere que los infantes humanos consideran "significativos" los actos de los demás, incluso cuando no entienden el punto de la acción y, tal suposición de significado, puede entenderse como un "acto de fe", en el sentido explicado anteriormente⁷⁸. El hecho de que los infantes imiten de esta manera nos puede dar buenas razones para pensar que el aprendizaje mediante la *imitación* se basa, no sólo en reconocer al otro como agente intencional, sino en la suposición de que la acción que se imita en sí misma puede poseer un sentido o un significado objetivo, aunque su interpretación sea ampliamente variable.

Ahora bien, considero que este planteamiento puede aplicarse plausiblemente al argumento del conocimiento moral general dotado evolutivamente.

Primero, es posible que dicho tipo de conocimiento moral pudo beneficiarse de la imitación (excesiva) de las estrategias de nuestros antecesores y como tal, de esas interpretaciones variables pero sujetas a un sentido o a un significado objetivo de la intencionalidad denotada en la acción; esto debido a que tal sentido o significado estuvo anclado al resultado plausible y exitoso de la acción, dando lugar así a las determinaciones generales que constituyeron las estrategias perniciosas o exitosas para el florecimiento y prosperidad de la especie.

⁷⁷ Para una discusión acerca de la distinción entre imitación y *mímesis*, donde la imitación refiere a réplicas exactas y *mímesis* a similitudes, ver Etzelmüller, 2016, p.198.

⁷⁸ Para las discusiones acerca de las concepciones de la imitación y la imitación excesiva, así como también su utilidad entre los diferentes escenarios de laboratorios y de la vida real, ver Etzelmüller (2016) p. 209-211.

Esto permite argumentar a favor de la existencia de aquellas orientaciones amplias y generales universalmente compartidas que constituyen el conocimiento moral general respecto a lo que clasificamos como correcto, incorrecto, etc, que, aunque sujetas a criterios objetivos, sus interpretaciones pueden ser ampliamente variables.

Serían objetivas debido a que provienen de aquellas estrategias exitosas que beneficiaron la prosperidad de la aptitud humana, pero permiten diversas formas de interpretación y por tanto diversas formas de implementación.

Con ello, es posible dar cabida a mi propuesta concerniente a que la moral es favorecida por aquel tipo de conocimiento moral general dotado evolutivamente que comprende las orientaciones normativas universalmente compartidas entre las poblaciones humanas; argumento otorgado por el computacionalismo. Pero para extraer los contenidos particulares de las situaciones concretas, he planteado que también requerimos del enactivismo, pues debido a su aspecto *situado* emergente de su metodología *relacional*, es capaz de hacerlo. Por consiguiente, aquí es preciso articular el planteamiento que se aproxime a una posible explicación pertinente a la evolución de dicho conocimiento moral general, el cual se acomodaría de la siguiente manera:

1. De la parte natural, primero, se vio que la condición intrínseca de vulnerabilidad del cuerpo humano fue un factor constituyente que dio lugar a la necesidad de cooperación en el hombre. Segundo, se consideró que es plausible admitir que al menos llegamos con una herencia evolutiva constituida por aquel núcleo biológico de rasgos heredables que incluye empatía e instintos sociales integrados con ciertos mecanismos que subyacen a nuestras orientaciones normativas generales provenientes de las estrategias exitosas y/o perniciosas de nuestros antepasados; lo que parece ser parte de nuestra dotación natural (*i.e.*, se consideró que llegamos instruidos en algún grado con ciertas predisposiciones naturales para desarrollar aquel tipo de conocimiento moral general favorecido por la evolución).
2. De los argumentos de la construcción de nicho, se vio que se constituyó un nicho cultural que propició la emergencia de las conductas pro-sociales humanas, el cual se articuló y heredó a partir del aprendizaje cultural, mismo que se benefició de la imitación, de la aloparentalidad, de la intersubjetividad y de la comunicación simbólica.

3. De ahí, se consideró que, a partir de la influencia recíproca gen-cultura que postula la genómica cultural, emerge una genuina interacción dinámica bio-cultural capaz de dar lugar a un tipo de conocimiento moral general, el cual se instaure a partir del núcleo biológico de rasgos heredables, pero también a partir del nicho dinámico favorecido por el aprendizaje por imitación, lo que en conjunto benefició tanto la selección de ciertos mecanismos que están detrás de aquellas orientaciones normativas incontrovertibles y universalmente compartidas, como también a la emergencia de su contenido general (nuestras nociones amplias de lo correcto/incorrecto, bueno/malo, etc.), el cual incluye una base *objetiva* pero a su vez soporta interpretaciones *multivalentes*.

Así, pues, si mis consideraciones son correctas, de los principios de la escuela de Santa Bárbara, se podría aceptar cierto tipo de arquitectura modular, pero una modularidad abierta e interactiva tal como la súper-modularidad, que se corresponde bien con la información pragmática, por lo que, nuestra dotación cognitiva vendría configurada con aquellas orientaciones normativas generales (o principios morales pre-lingüísticos), pero a la vez sería capaz de interactuar de manera sustancial con las influencias del contexto y del entorno.

Y, concerniente a la hipótesis de la construcción de nicho, se hizo ver la notable relación que prevalece entre el aprendizaje social y el nicho cultural humano, en donde se revisaron sus condiciones, sus características y de ello se destacó que la imitación fue un elemento decisivo para el desarrollo del aprendizaje cultural y como tal dio lugar a la herencia ecológica que propició la emergencia de conductas prosociales y de aquellas orientaciones normativas generales que guían nuestros cursos de acción.

De ahí, pienso que arribamos a una posible explicación (de *grano grueso*) acerca de la presencia del tipo de conocimiento moral general que es parcialmente dotado por la evolución y que considera tanto la evolución natural como la evolución cultural.

Tras ver el argumento evolutivo, ahora revisemos la cognición computacionalista-enactiva que defiende.

4.4. Cognición computacionalista-enactiva

Este apartado se enfoca en esbozar cómo funcionaría la parte computacionalista y la parte enactiva, respectivamente, en la arquitectura cognitiva PI.

Al respecto, es pertinente mencionar que mi pretensión aquí es modesta, debido a que únicamente intento aproximarme a una posible propuesta que permita articular coherentemente cómo trabajarían los dos modos cognitivos a fin de conciliar las virtudes que ellos ofrecen, pues una descripción detallada requerirá de mayor investigación mediante un trabajo multidisciplinar y escrutinio empírico, lo que considero, puede abrir nuevas rutas para trabajos ulteriores.

Como punto de partida, recapitulemos cómo funciona el modelo cognitivo de cada marco conceptual, para después, esbozar cómo se corresponderían ellos con la arquitectura PI.

1. Computacionalismo:

Percepción-cognición (procesamiento de información)-acción

2. Enactivismo:

Percepción- $\leftarrow\rightarrow$ acción

A partir de ello, propongo que (1) se corresponde con la parte de PI que soporta los procesamientos de información representacionales y guiados por principios. Y, (2) se corresponde con la parte de PI que acepta que parte de su diseño sea aprendido mediante la experiencia que surge de la acción y de la interacción continua con el ambiente.

La parte (1) está plenamente constituida por la arquitectura computacionalista, por lo que se ajusta plausiblemente con la súper-modularidad y por consiguiente con su naturaleza representacional, abierta e interactiva.

Concerniente a (2), lo relevante es notar que, mediante la *acción* directa con el entorno (sin la mediación de procesamientos de información representacionales), es capaz de soportar el aprendizaje, el cual puede ser desde cero. Recordemos que aquí la entidad teórica clave es la *acción*, debido a que es a partir de ella que el sistema aprende y experimenta directamente su medio, arribando a patrones aprendidos bien estructurados que le permiten más acciones para navegar y cumplir exitosamente las tareas relevantes, lo que es compatible con el enactivismo.

Asimismo, es plausible pensar que en esta arquitectura se contaría con la intervención tanto de las heurísticas rápidas y frugales como de la cognición *ultrarrápida*, cada una correspondiéndose con sus elementos habilitadores.

De esta manera, entonces, propongo que tendríamos una arquitectura cognitiva capaz de soportar los principios, las metodologías y las entidades núcleo que cada marco postula, en donde se puede hablar coherentemente de representaciones y principios que guían la cognición, pero también, de la emergencia de ésta a partir de la acción e interacción con el ambiente y de su aspecto genuinamente relacional.

Concerniente a las incidencias concretas que pueden ocurrir (o no) entre ellos y cómo éstas funcionarían, es tema que requiere de mayor escrutinio. No obstante, y con esto en mente, revisemos cómo se correspondería esto con la cognición social y con la cognición moral, para después determinar aproximadamente qué papel jugaría el computacionalismo y el enactivismo en la cognición socio-moral, según mi propuesta.

4.5. Cognición social computacionalista-enactiva

Dada la arquitectura cognitiva PI, veamos cómo quedaría el modelo de la cognición socio-moral computacionalista-enactiva, la cual es capaz de soportar las perspectivas epistemológicas de 2ª y de 3ª persona, así como también, de beneficiarse de las virtudes de ambos modelos cognitivos. Revisemos esto con más detalle.

Comienzo instanciando el modelo socio-cognitivo computacionalista y enactivo, respectivamente:

Modelo socio-cognitivo computacionalista:

El sujeto adscriptor *observa* al otro y percibe la información relevante (señales del entorno y conducta manifiesta) → procesa tal información mediante el proceso de atribución mental postulado (ya sea TT o TS o una hibridación de ellos), lo que le permite producir una representación del otro y de sus estados intencionales → obtiene un estado resultante, el cual atribuye al otro. Con ello, se explica y predice la conducta del otro. Respecto a la acción social, es la representación del otro y de sus estados intencionales lo que le permite al sujeto adscriptor planificar cómo actuar de una manera socialmente apropiada, ante lo cual, responde ejecutando ese plan con comandos motores lo mejor posible.

Modelo socio-cognitivo enactivo:

En una interacción (digamos, diádica), ambos sujetos se perciben directamente y actúan en consecuencia (hay un acoplamiento sensoriomotor entre ellos en donde sus acciones son guiadas perceptualmente), lo que resulta en que ambos participan en una mutualidad

directa que da lugar a participar en la creación de nuevos significados propios de esa interacción específica y por tanto ambos responden a exigencias concretas de la situación. Con ello, ambos interactuantes participan y responden a la mutualidad que la interacción en sí misma exige.

De ello, desde un mapeo de *grano grueso*, se podría establecer que todas aquellas tareas socio-cognitivas que emplean la rutina de atribución mental, corresponderían a la parte (1) de la arquitectura PI, mientras que aquellas habilitadas por la perspectiva de 2ª persona y beneficiadas por capacidades corporizadas más básicas, tales como la percepción directa, corresponderían a la parte (2) de PI.

Con ello, podríamos contar con una arquitectura cognitiva integral capaz de utilizar estrategias alternativas, según requiera el caso. Esto permite considerar casos generales, pero también casos particulares, lo que posibilita el análisis de una gran gama de situaciones sociales, desde las cotidianas y de fácil manejo hasta las problemáticas, extrañas o ambiguas, entre otras.

Así, en ocasiones emplearíamos rutinas de atribución mental beneficiadas por la *observación*, y en otras, *participaríamos* directamente con el otro, nos involucraríamos emocionalmente con él y percibiríamos inmediatamente sus estados. Esto conduce a pensar que las estrategias socio-cognitivas disponibles pueden trabajar independientemente o co-participando entre ellas; lo que presentaría ventajas ante las estrategias excluyentes.

Ahora, veamos cómo se beneficiaría la cognición moral de una arquitectura integral así definida, tal como PI.

4.6. Cognición moral computacionalista-enactiva

Una vez que ya se ha esbozado cómo sería el modelo integral socio-cognitivo computacionalista-enactivo, ahora corresponde hacer explícito cómo funcionaría aproximadamente el modelo concerniente a la cognición moral.

Para ello, en primer lugar, recapitulemos cómo trabaja el modelo computacionalista y el enactivo, respectivamente:

Cognición moral computacionalista:

Es habilitada por la arquitectura súper-modular, que es interactiva, abierta, multidireccional y se beneficia de razonamientos heurísticos, con lo que involucra de manera sustancial la información pragmática. Los principales procesamientos de información cargados de contenido que dan lugar a la cognición moral desde este punto de vista son: ToM, entendimiento de intenciones, entendimiento de emociones, principios morales pre-lingüísticos y empatía. La interacción entre ellos es lo que habilita la competencia moral humana.

Cognición moral enactiva:

Es habilitada por la perspectiva de 2ª persona que otorga un papel de *participante* a los agentes que interactúan; característica que implica responsabilidad moral. Y, consecuentemente, se beneficia de al menos tres aspectos fundamentales: 1) de un *tirón* moral fenomenológico sentido en el cuerpo vivo, 2) de la normatividad co-regulada que emerge de la interacción, y, 3) es heterónomamente regida por el contexto socio-cultural.

En conjunto, estos tres niveles (o aspectos) dan lugar a la cognición moral enactiva, la cual se encargaría (desde un mapeo de *grano grueso*) de la acción moral pre-reflexiva que surge del confrontamiento *directo* con el otro y con el mundo (Varela, 1999), proceso que es habilitado por bucles de percepción-acción. Y, por otra parte, la cognición moral enactiva intervendría crucialmente en los juicios morales efectuados desde la perspectiva de 2ª persona. Este tipo de juicios morales son compatibles con aquellos postulados por Strawson (1974). Más adelante desarrollo este tema.

Ahora bien, en este punto, se deben resaltar las diferencias de fondo entre ambos modelos cognitivos: si bien, el computacionalismo también habla del infante como *participante* en su proceso de aprendizaje (en relación con su entorno y sus interacciones sociales), se debe destacar que, en el enactivismo, que preeminentemente otorga un papel de *participante* al agente, lo efectúa desde su metodología relacional, en donde el agente y el entorno se co-habilitan continuamente en las interacciones.

Esto marca una diferencia sustancial, pues, en primer lugar, como se ha hecho ver, cada modelo cognitivo se enfoca en distintos aspectos del fenómeno a estudiar y como tal es habilitado por diferentes elementos que le subyacen. En segundo lugar, siguiendo a van Grunsvan (2018), he propuesto reconocer que las interacciones corporizadas en segunda persona o, más bien, la *experiencia* de los involucramientos en segunda persona da lugar

a nuestra orientación de ser *participantes* en las interacciones que sostenemos, lo que implica tener responsabilidad por el otro. En otras palabras, se ha argumentado que la perspectiva experiencial de las interacciones en segunda persona implica una orientación de participación, lo que a su vez implica responsabilidad hacia el otro. Bajo esta lectura, los participantes de cada interacción tienen grados de responsabilidad por el otro, movimiento que representa una diferencia importante entre los modelos.

De la misma manera, tómesese en cuenta que cada modo cognitivo apela a recursos distintos respecto al *saber-cómo* moral: el computacionalismo recurre al uso de heurísticas rápidas y frugales que se benefician del uso de reglas para la toma de decisiones en un sinnúmero de escenarios moralmente relevantes, mientras que el enactivismo apela a la *cognición ultrarrápida*, la cual toma en cuenta las contribuciones del agente y del contexto en el desempeño cognitivo (Järvilehto, 1998), con lo que pone en marcha los esquemas sensoriomotores del agente para arribar a la mejor opción disponible en el contexto.

Así, pues, tendríamos un modelo integral de la cognición moral, el cual se beneficiaría fundamentalmente de diversas estrategias, tal como la toma de perspectiva, el uso de ToM, el entendimiento intencional, el uso de representaciones, etc., pero también, de estrategias corporizadas más básicas, las cuales favorecen, al menos, la acción pre-reflexiva y los juicios morales en 2ª persona. Así, contaríamos con un modelo que puede trabajar en diversos niveles.

No obstante, frente a ello, cabría objetar acerca de algunos problemas que pueden surgir o tensiones que pueden hacer colapsar la propuesta, por lo que a continuación las abordo directamente.

4.7. Algunos problemas que pueden surgir

Primero, si la súper-modularidad es capaz de explicar lo indeterminado de cada contexto, ¿por qué recurrir al enactivismo? La respuesta radica en que en mi propuesta otorgo autonomía a los procesos supra-personales, lo que conduce a tomar en cuenta el aspecto relacional en el estudio de la cognición moral. Al respecto, defiendo que esta dimensión debe ganar un lugar en su estudio, pues ofrece un análisis estricto de los aspectos y factores interactivos, *i.e.* de aquello que emerge de las interacciones entre individuos y de las interacciones entre el agente y el entorno, de lo que surgen significados y normatividades en un dominio cognitivo irreductible al nivel individual. Y, dicho análisis

efectuado desde una dimensión relacional es posible gracias a los principios enactivos, que parten en origen de una metodología relacional.

Así, pues, como se hizo ver, definiendo la complementación de ambas metodologías (computacional y enactiva) para dar cuenta de la cognición socio-moral.

De la misma manera, cabe preguntar, si el NMI se postula como un saber-cómo moral (con su correspondiente saber-qué), ¿esto no colapsaría con la acción pre-reflexiva del enactivismo? Mi respuesta es que no y la razón es que en mi propuesta la acción pre-reflexiva enactiva emerge fundamentalmente de la metodología relacional de la interacción, la cual es favorecida por los bucles de percepción-acción que no suponen separación modal entre ellas.

Para ser clara, aquí recurramos otra vez a la diferencia entre los elementos que habilitan a cada modo cognitivo respecto al saber-cómo moral: el computacional funciona mediante heurísticas rápidas y frugales y el enactivo mediante la cognición ultrarrápida, lo que deriva en que la primera es guiada por el uso de reglas y la segunda por la percepción, y como he mencionado, hay evidencia empírica que podría avalar la existencia de ambos modos cognitivos (automáticos), tales como la percepción y clasificación de actos dañinos (ver Gray, 2018). La diferencia radicaría en el aspecto *relacional* del enactivismo, pues en ella, se involucra la percepción de la situación, las posibilidades de acción que ella ofrece, las habilidades y/o competencias del agente participante y la normatividad co-habilitada que de ello surge, de lo que deriva la intervención del nivel supra-personal (autónomo) en los procesamientos cognitivos socio-morales.

En este punto, es preciso aclarar que en el enactivismo se consideró el aspecto relacional de los valores morales, es decir, ellos no radican únicamente en el sujeto o afuera en el mundo (natural o social), sino que emergen *relacionalmente* de la interacción entre estos, incluso se hizo ver que ellos son negociados momento a momento entre los participantes de la interacción. En cambio, una consecuencia del computacionalismo es que primero percibimos la situación y posteriormente hay un mecanismo seleccionado encargado de asignarle un valor (ver Barrett, 2015), por lo que aquí se vuelve relevante la separación modal percepción-acción sostenida por el computacionalismo.

En otras palabras, el NMI explica contundentemente la competencia moral como un saber-cómo, pero éste se corresponde con los principios computacionalistas, lo que deriva en

apelar a mecanismos basados-en-principios encargados de la asignación de valores, en su lugar, en el enactivismo, el valor surge *relacionalmente* de la interacción misma, por lo que la percepción-acción es directa, de tal suerte que habría una respuesta moral directa e inmediata que favorece la acción pre-reflexiva enactiva.

En mi interpretación, ambas estrategias forman parte de la cognición moral competente, en donde la estrategia que se emplea obedece a lo que cada situación demanda. Considero que ello ofrece un crisol cognitivo más abarcador capaz de posibilitar diversas interpretaciones, acciones e interacciones que componen el amplio espectro de la cognición socio-moral.

Así, pues, pienso que la posibilidad de contar con diversas estrategias presenta ventajas frente a otras explicaciones de carácter excluyente, puesto que, de esta manera, se podría cumplir con una amplia gama de situaciones morales habilitadas por diferentes recursos. Esto nos lleva al siguiente tema que es la normatividad moral computacionalista-enactiva, en el que intento hacer ver lo benéfico que resultaría fusionar las virtudes de ambas posturas.

4.8. Normatividad moral computacionalista-enactiva

A partir de los postulados normativos de cada marco conceptual, tenemos que el computacionalismo brinda el argumento a favor de un tipo de conocimiento moral general dotado evolutivamente, lo que asegura la objetividad moral, y, el enactivismo, debido a que surge de la interacción y se instituye como práctica social, es capaz de explicar el aspecto *situado* de la normatividad moral.

Arguyo que considerar estos dos aspectos en conjunto permite postular una herramienta evaluativa capaz de generar juicios situados pero anclados a criterios objetivos.

El argumento es como sigue: el conocimiento moral general se constituye fundamentalmente de los principios morales prelingüísticos que se corresponden con aquellas orientaciones normativas incontrovertidas y universalmente compartidas. La existencia de algunos de estos principios ha sido mostrada empíricamente, tales como: ayudar a otros es bueno, el sufrimiento es malo, obstaculizar a otros es malo y cada quien tiene lo que se merece. Ellos serían la *condición de posibilidad* de más valores morales y sus derivaciones se corresponderían con el contexto y la cultura.

Por su parte, desde el enactivismo se propone un principio moral básico con valor transcultural, a saber, *el cuidado de las autonomías*. No obstante, la manera en cómo se efectúa este principio es dictado fundamentalmente por el contexto y la cultura, lo que da lugar al aspecto situado de la normatividad.

Arguyo que el principio enactivo concerniente al *cuidado de las autonomías* puede sumarse plausiblemente (*prima facie*) a aquellas orientaciones incontrovertidas y universalmente compartidas que forman parte de nuestro aparato cognitivo, pues por su generalidad, se ajusta bien con ellos, y, por ende, con el tipo de conocimiento moral general postulado.

Así, tenemos que dicho conocimiento moral general guía normativamente nuestros cursos de acción, pero también, ellos son regidos y dictados por el contexto y la cultura.

Ahora bien, en la parte computacionalista se planteó que los principios morales prelingüísticos dotados evolutivamente tienen la capacidad de ser realizables de múltiples maneras, lo que significa que sus derivaciones obedecen y se ajustan a las exigencias del contexto en el que toman lugar.

De esto, veo dos puntos de compatibilidad potenciales entre el computacionalismo y el enactivismo: 1) el cuidado de las autonomías plausiblemente puede sumarse a los principios morales prelingüísticos que postula el computacionalismo, y; 2) tales principios, al igual que el cuidado de las autonomías, poseen pluralismo de realizabilidad.

Estas compatibilidades son valiosas porque son las que permiten plantear dicha herramienta evaluativa capaz de generar juicios situados pero anclados a criterios objetivos.

Veamos cómo funcionaría en un ejemplo contundente. En el capítulo 2 expuse un dilema moral paradigmático: la ablación femenina, el cual pienso que puede ser abordado por dicha herramienta de la siguiente manera:

Primero, el enactivismo ofrece una útil distinción entre *carácter* moral y *contenido* moral, lo que permite distinguir acciones y prácticas opresivas de las que no lo son, es decir, de aquellas que en su lugar potencializan las autonomías de los agentes.

Si efectuamos un análisis de grado y situado de dicha práctica, y que además tome en cuenta tal distinción, podemos obtener diversos juicios. Esto se logra por analizar los motivos y creencias que cada práctica particular detenta, en donde se toma en cuenta la

voz de las agentes implicadas como también las circunstancias específicas en las que toma lugar la práctica.

Al respecto, en co-autoría con mi colega Patricia Melchor (en revisión), proponemos el ejemplo de diversas comunidades que practican la ablación femenina, pero cada una de ellas sustentadas por distintas motivaciones y creencias. Consideramos que a partir de un análisis de grado y de un escrutinio detallado que emplea la distinción enactiva de contenido moral y carácter moral, es posible arribar o no a su justificación moral.

En general, la práctica de la ablación femenina, que en la actualidad se realiza mayormente en comunidades africanas, ha llamado la atención global y ha generado la reprobación total de su ejecución y de sus consecuencias. No obstante, debe tomarse en cuenta que los discursos justificatorios para abolir tal práctica se basan en concepciones muy generalizadas de sus practicantes, es decir, se les concibe como entes limitadas e inertes que sólo son receptoras pasivas de la mutilación, sin considerar que tal práctica puede obedecer a distintas causas y detentar diversos sistemas de creencias, que incluso, sean moralmente justificadas (Young Eun Nam, 2018).

En torno a ello, proponemos que la naturaleza situada de la propuesta normativa enactiva puede entrar en juego y ofrecer un análisis de grado que vaya más allá de los juicios generales, universalistas y abstractos que no aceptan excepciones y reprueban enteramente dicha práctica, pues hay casos en los que, incluso, se ejecuta por la decisión autónoma de las agentes.

Por supuesto, los casos que implican la degradación de las autonomías de las agentes que resultan en daños físicos y psicológicos irrevocables, sin duda alguna, son prácticas de *carácter moral* incorrecto y con *contenido moral* incorrecto, puesto que, a la base de la mutilación misma, se encuentran creencias en detrimento de la mujer (Kouba y Muasher, 1985, p. 103).

No obstante, existen casos en los que la ablación femenina forma parte de un ritual, por ejemplo, el ritual *irua ria atumia*, que equivale a un proceso de iniciación similar al de los hombres de la comunidad, el *irua ria anake*, y es más similar a una circuncisión que a las mutilaciones altamente intrusivas que asociamos con la ablación femenina. Este caso es analizado por Wairimū Ngarūiya (2011), quien hace ver que en esta comunidad keniana llamada *Gĩkũyũ*, estos procedimientos (tanto femeninos como masculinos) preparaban

física y psicológicamente a los participantes, fortaleciendo y dando un sentido de pertenencia que incluía una serie de actividades de formación, iniciación e integración a las prácticas adultas de la comunidad, en donde se daba a los iniciados información acerca del funcionamiento del grupo y se practicaban rituales que generaban una fuerte cohesión entre ellos (ver Pescador y Melchor, en revisión, pp.32-33).

De esta manera, Ngarūiya tiene como objetivo dejar de lado los juicios generalizados y en su lugar ofrecer un análisis integral capaz de tomar en cuenta las voces de los participantes y denotar la complejidad de las prácticas, lo que deriva en significarlas de manera distinta a lo que se ha pensado desde una visión parcial, con lo que, incluso, serían moralmente justificadas; pues en estos casos, la ablación es resultado de la decisión autónoma de las mujeres.

Así, “en estos casos específicos, la práctica tiene un claro *carácter moral*, en tanto implica un acto que cuida y aumenta las posibilidades de acción e interacción que contribuyen a afirmar la vida autónoma de las mujeres en esas comunidades. Su *contenido moral* sería considerado correcto, debido que a su base se encuentran creencias e idiosincrasias que favorecen el cuidado de las autonomías, tanto masculinas como femeninas. Sin embargo, tal evaluación sólo puede darse en este caso concreto [...]” (*Ibid.*, p.34).

Ahora bien, para mis propósitos, es importante señalar que el aspecto *situado* que ofrece la normatividad enactiva, permite este tipo de juicios que en ocasiones podrían justificar moralmente casos controversiales, sin embargo, siempre se requerirá una base imparcial anclada a criterios objetivos que eviten la apropiación cultural de los fundamentos morales, lo que es otorgado plausiblemente por el conocimiento moral general dotado evolutivamente, o al menos así lo defenderé.

De esta manera, considero que la fusión de la normatividad computacionalista y enactiva posibilitaría una útil herramienta evaluativa capaz de generar juicios *situados* pero anclados a criterios *objetivos*, y que tal vez pueda invitar a nuevos acercamientos evaluativos respecto a dilemas morales.

Con esto en mente, ahora abordo el esquema del modelo cognitivo computacionalista-enactivo que habilitaría este tipo de juicios.

4.8.1. Esquema del modelo cognitivo de la normatividad computacionalista-enactiva

De lo anterior, propongo que el modelo cognitivo de la normatividad computacionalista-enactiva que planteo quedaría instanciado en el siguiente esquema:

En una matriz relacional⁷⁹ como la siguiente (ver Fig. 2), la “flecha” marcaría lo universal (el conocimiento moral general dotado evolutivamente) y la “espiral” que le envuelve, marcaría los cambios de lo individual, lo inter-individual y lo supra-personal, siempre moldeando a aquellos criterios objetivos desde un aspecto relacional. Es decir, lo *objetivo* siempre permanece y se mantiene, pero está constantemente influenciado por los factores *situados* supra-personales, de tal manera que ambos rasgos se confluyen continuamente.

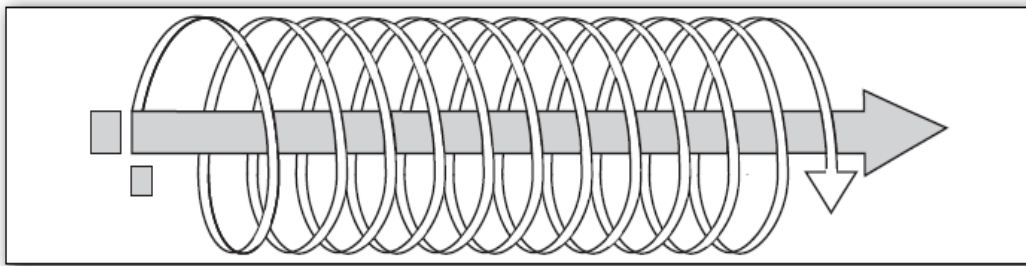


Fig.2. Esquema extraído de Gutiérrez-Martínez, F. (2005). *Teorías del desarrollo cognitivo*. McGraw-Hill, España, p.333.

De ello, propongo que tendríamos una herramienta evaluativa capaz de generar juicios anclados a criterios objetivos, los cuales permanecen y se mantienen, pero a su vez, correspondiéndose con el carácter *situado* de cada caso desde su aspecto relacional de valores.

Ahora corresponde ver la co-participación de los elementos presentados.

4.9. Co-participación de los elementos presentados

Aquí es muy importante tener en cuenta que las pretensiones del enactivismo son moderadas, y por ello, este enfoque únicamente busca complementar las explicaciones disponibles. En esta línea, es plausible pensar en que el enactivismo tiene el potencial de efectuar algunas tareas socio-morales, no todas las tareas.

⁷⁹ La matriz relacional es propuesta por Overton (1998) y refiere a una estructura capaz de soportar lo objetivo que es una característica que siempre se mantiene estable, pero a su vez, es capaz de soportar los cambios que el contexto y la situación demanden.

Se vio que el lado computacionalista es habilitado por la arquitectura súper-modular, en donde los juicios morales y las conductas morales son efectuados mediante la interacción de diversos procesamientos de información cargados de contenido, tales como ToM, entendimiento intencional, entendimiento de emociones, principios morales pre-lingüísticos y empatía, tal como se definió en el capítulo 2.

Ahora bien, la literatura de la psicología moral lo corrobora al otorgar ejemplos que ilustran el uso de este tipo de cogniciones morales compatibles con el computacionalismo (ver capítulo 2), esto debido a que los estudios de la psicología moral se han dirigido a los estados mentales moralmente relevantes y con ello a la economía cognitiva del agente moral. No obstante, esta disciplina se ha centrado exclusivamente en los aspectos mecánicos, en lugar de experienciales, de la percepción moral, del pensamiento moral, de la deliberación moral y de la emoción moral (ver Kriegel, 2007). Asimismo, ha intentado rastrear qué función desempeñan los estados mentales relacionados al ámbito moral en la economía mental del sujeto, pero ha ignorado lo que es para él experimentarlos.

Por lo anterior, a continuación, desde un mapeo de *grano grueso*, me enfoco en plantear las tareas de las que se encargaría el enactivismo en mi propuesta, pues pienso que éste es capaz de abordar ese lado experiencial no visto, y también, las intervenciones *relacionales* del nivel supra-personal, es decir, definiendo que el enactivismo es capaz de capturar las dimensiones fenomenológicas, interactivas y corporizadas del fenómeno a estudiar.

En torno a ello, sostengo que el computacionalismo es contundente al explicar la cognición socio-moral (se encarga de la mayoría de las tareas socio-morales), de hecho, he defendido que sin su explicación no podríamos tener una comprensión completa de la misma, no obstante, sustento que el enactivismo se encargaría, al menos, de dos tareas socio-morales: de la acción pre-reflexiva guiada perceptualmente y habilitada por la cognición ultrarrápida, y de algunos juicios morales en 2ª persona. Veamos esto con más detalle.

- **Enactivismo: acción moral como un saber-cómo pre-reflexivo y juicios morales en 2ª persona**

Siguiendo la coherencia conceptual enactiva, queda de manifiesto que, debido a sus principios y a su modelo cognitivo (el cual no admite separación modal entre la percepción y la acción) es enteramente capaz de explicar la acción moral pre-reflexiva que es guiada perceptualmente y como aquello que emerge del confrontamiento directo con el mundo, sin

desacoplamientos ni mediadores representacionales. Por ello, he considerado que podemos hablar de dos sustratos habilitadores de las competencias pertinentes al saber-cómo moral: uno habilitado mediante razonamiento heurístico y el otro guiado perceptualmente, el cual deriva naturalmente de la concepción enactiva.

A partir de la evidencia empírica presentada en el capítulo 3, podemos comenzar a hablar del segundo de ellos, pues se explicó cómo respondemos directamente al otro, cómo nos acoplamos y respondemos a la estructura de la interacción y de la situación y, siguiendo esta línea, cómo respondemos a la situación moralmente relevante y a las necesidades concretas que ésta exige mediante bucles de percepción-acción.

En ello, estarían involucrados factores corporizados tales como la perspectiva de 2ª persona, la percepción directa, el involucramiento emocional, la intencionalidad compartida, entre otros, todos ellos emergentes de los acoplamientos sensorio-motores en marcha y de los continuos bucles de percepción y acción que de ello surgen.

Así, la acción moral pre-reflexiva enactiva se explicaría coherentemente a partir de los argumentos y de la evidencia presentada, en donde diversos ejemplos podrían entrar en juego. A estos fines, es pertinente recordar, además, que van Grunsven (2015) sostiene que nuestra orientación perceptiva de los demás es parte integral de cómo interactuamos con ellos: cómo fallamos y cómo logramos responderles adecuadamente como personas.

Concretamente, un ejemplo paradigma que propone Varela respecto a la acción moral pre-reflexiva, es que, en cuanto percibimos que ocurre un accidente cerca de nosotros, acudimos *inmediatamente* a ayudar, en donde nuestra acción es provocada inmediatamente por la percepción de la situación⁸⁰. En esta línea, considere también las situaciones (de charla) en las que percibimos la incomodidad de alguien por el tema de conversación abordado y actuamos (*inmediatamente*) intentando cambiar el tema para aliviar su malestar (Varela, 1999). Según Varela, estos son ejemplos en donde no hay

⁸⁰ En sentido concomitante, Jayawickreme y Chemero (2008) ofrecen valiosos ejemplos de ello, en donde la situación percibida conduce a responder (pre-reflexivamente) en consecuencia. La ética del cuidado también ofrece un amplio cuerpo de evidencia empírica enfocado en sostener cómo a partir de estos fundamentos corporizados respondemos a las necesidades concretas y moralmente relevantes de *otro* particular (p. ej. ver Kittay, 2009). Asimismo, van Grunsven (2015) ofrece un tratamiento robusto respecto a la percepción enactiva y su relación con aspectos morales.

desacoplamientos ni mediadores representacionales en nuestras acciones, sino más bien, surgen de la percepción directa de la situación.

No obstante, en este punto alguien podría cuestionar si estas reacciones son habilitadas por la percepción y por la cognición ultrarrápida, o bien, si son basadas en principios y habilitadas por el razonamiento heurístico o, incluso, si son *gatilladas* por la coparticipación de ambos. La respuesta específica a ello implicará un profundo estudio empírico, trabajo que posibilitaría demarcar la genuina intervención de ambos modos cognitivos y con ello el papel del aspecto *relacional* en la acción moral. Sin embargo, lo que se puede extraer a partir de la evidencia empírica presentada es que ambos modos cognitivos se legitiman, cada uno enfocado en aspectos distintos.

Otra cuestión a considerar es si el enactivismo también recurriría a los principios morales pre-lingüísticos. Dado el marco integrativo que propongo, facultado por la arquitectura PI, es posible hablar de una co-participación entre los elementos aludidos, en donde sería posible conceder que el enactivismo es guiado por principios, ante lo cual no es necesario apelar a la articulación lingüística de los mismos, sino al habilitamiento de un repertorio de esquemas sensoriomotores adquirido a lo largo de la historia de interacciones del agente (de Souza, 2013). En mi interpretación, tal esquema sería coadyuvado por la arquitectura cognitiva dotada evolutivamente que alberga las propensiones naturales para el desarrollo del conocimiento moral general basado en aquellos principios morales pre-lingüísticos.

Un rasgo importante que puede prestar apoyo a ello es que, según Gray (2018), la gama de la percepción de daño es parte de la mente humana y ésta es claramente escalonada, pues los participantes presentan los mismos tiempos de reacción en las tareas relevantes: “La naturaleza gradual del daño también existe en la mente de las personas. Nuestro trabajo revela que cuando las personas hacen juicios rápidos sobre la nocividad de diversos actos inmorales, que van desde el asesinato hasta el chisme, sus tiempos de reacción forman un claro gradiente (Schein y Gray, 2015; Estudio 4).” (ver Gray, 2018, p. 42).

La clasificación de actos dañinos es la siguiente: el asesinato y la violación se clasifican como dañinos en 600 ms, seguidos por el adulterio y el abuso en 700 ms, la prostitución y la mentira en 800 ms y finalmente la falta de respeto y los chismes en 850 ms. (Gray, 2018, p.42). Cabe destacar que estos actos son juzgados como inmorales en los mismos tiempos de reacción. En particular, estos datos son relevantes debido a que robustecen la idea de que la arquitectura cognitiva humana cuenta con propensiones naturales para desarrollar

las nociones normativas morales generales y universalmente compartidas, pues, al parecer, contamos con la misma arquitectura cognitiva que guía nuestras diversas clasificaciones de actos dañinos.

Sin embargo, determinar específicamente en qué tarea interviene cada modo cognitivo, requerirá más trabajo y escrutinio empírico, no obstante, aquí he propuesto que la mayor parte del trabajo lo hace el computacionalismo, pero que existe un espacio para el enactivismo.

En esta línea, lo que resta, entonces, es explicar los juicios morales en 2ª persona, tema que presento a continuación.

- 2ª persona, participante y responsabilidad moral

Primero, como se vio, Reddy propone una poderosa liga entre la experiencia con el otro y una experiencia ética que “me obliga a responder”⁸¹. De ello, en la psicología del desarrollo comienza a trazarse un vínculo directo entre la noción de *participante* y la noción de responsabilidad moral desde la perspectiva de 2ª persona.

Esto encuentra apoyo teórico. En particular, van Grunsven (2018) nutre estos argumentos al asociarlos con la ética del cuidado y con la propuesta de responsabilidad moral de P.F. Strawson (1974); explicaciones que ponen en primer plano las interacciones en 2ª persona y su afectividad intrínseca.

Como mencioné, van Grunsven (2018) sostiene que las interacciones corporizadas en segunda persona o más bien la *experiencia* de los involucramientos en segunda persona dan lugar a nuestra orientación de ser *participantes* en las interacciones que sostenemos, lo que, según la autora, implica tener responsabilidad por el otro. Es decir, la perspectiva *experiencial* de las interacciones en segunda persona implica una orientación de participación, a partir de lo cual, siguiendo a Strawson y a la ética del cuidado, Grunsven ha sugerido que dicha noción de participación implica responsabilidad hacia el otro. Así, desde este punto de vista, los participantes de cada interacción, tendrían responsabilidad por el otro.

Ahora bien, en línea con la ética del cuidado y con su perspectiva relacional, se arguye que la responsabilidad equivale a ejercitar nuestras capacidades para responder

⁸¹ Lo que es enteramente compatible con el *tirón moral* fenomenológico del enactivismo.

adecuadamente a las necesidades particulares del otro en la interacción (van Grunsven, 2018, p.143); es decir, en la interacción existe la posibilidad de responder apropiadamente al otro, o bien, de fallar y no responder a sus *demandas particulares*. Desde este punto de vista, hay grados de responsabilidad implícitos en las interacciones afectivas que sostenemos con otros agentes humanos.

Por su parte, Strawson, en su ensayo preeminente "Freedom and Resentment" (1974), sostiene que en el campo moral existe una profunda diferencia entre las perspectivas de 2a y 3a persona. Según Strawson, la responsabilidad moral tiene lugar desde una perspectiva de 2a persona en tanto que: 1) entiende las interacciones sociales como involucramientos afectivos y en ellos se prioriza la perspectiva de 2a persona; 2) la perspectiva de 2a persona posibilita un rango de actitudes y sentimientos que no estarían disponibles en una perspectiva de 3a persona, y; 3) en las interacciones sociales (así definidas) respondemos afectivamente al otro. Strawson llama a estas reacciones afectivas "actitudes reactivas". Entonces, el argumento es que, a partir de la perspectiva de 2a persona, reaccionamos *directamente* a las demandas del otro y de la situación, lo que da lugar a nuestras actitudes reactivas para adscribir responsabilidad a los agentes o, dicho de otro modo, para culparlos, elogiarlos, etc. (van Grunsven, 2015, 2018).

Por otra parte, de acuerdo con Strawson, tomar una perspectiva de 3a persona, *i.e.*, de mero observador, significa un modo extraño de abordar las implicaciones morales que la situación concreta exige, puesto que nos haría *suspender* nuestras actitudes reactivas. La perspectiva de 3ª persona de observación y predicción apunta a un enfoque distante de las personas, en el que *suspendemos* o *pausamos* nuestras demandas y expectativas cotidianas y, en cambio, consideramos al otro como "un objeto social". Es desde estos enfoques que van Grunsven (2018) comienza a trazar una explicación que otorga grados de responsabilidad moral a la orientación de ser participantes en la interacción.

- **Juicios morales en 2ª persona**

El enfoque en segunda persona que Reddy propone, evita tanto la incapacidad de la TS para comprender a los otros realmente como otros, como también, la debilidad de la TT por ignorar el papel que juega la experiencia afectiva en nuestro entendimiento de los demás.

Por supuesto, descubrir las intenciones, creencias y deseos de los demás y explicar y predecir su comportamiento es muy importante para la regulación de la vida social, sin embargo, los teóricos enactivos han cuestionado ampliamente si esta postura de

observación y predicción en tercera persona es tan fundamental y omnipresente como lo han asumido las teorías computacionalistas de la cognición social.

Como se vio, la evidencia empírica del Capítulo 3 apoya lo que sostiene Gallagher (2007), entre otros autores: que gran parte de nuestra vida social, ya sea como adultos o como agentes sociales en desarrollo, se negocia a través de interacciones en 2ª persona, las cuales son facilitadas por un entendimiento perceptivo directo de la conducta corporal expresiva de los demás, lo que permite una gama de posibilidades más o menos adecuadas de acción y respuesta (Gallagher, 2007, p. 354). Asimismo, como se vio, la propuesta de Grunsven (2018) otorga grados de responsabilidad a las interacciones corporizadas en 2ª persona, lo que contribuye a este punto de vista y es sustentado en la psicología del desarrollo por Reddy, quien defiende el surgimiento de una poderosa liga ética en las interacciones de 2ª persona.

Ahora bien, Strawson sugiere que existe una profunda diferencia entre un entendimiento perceptual en 2ª persona y una visión observadora de los demás en 3ª persona, lo que no significa que niegue, al igual que Gallagher, que la postura observacional u "objetiva" también está a nuestra disposición en nuestras relaciones sociales. No obstante, Strawson insiste en que no podemos adoptar esta perspectiva de observación en 3ª persona por mucho tiempo, ya que requiere de una *suspensión* de un rango de actitudes y sentimientos disponibles en una perspectiva de 2ª persona que, según Strawson, "forman parte esencial de la vida moral como la conocemos y se oponen a la objetividad de la actitud" (Strawson, 2008, p. 34), agregando que, "cómo es realmente estar involucrado en relaciones interpersonales ordinarias repercute en nuestro sentido moral" (Strawson, 2008, p. 23).

La idea de que nuestra obligación y responsabilidad ante los demás está íntimamente ligada a nuestra capacidad de verlos y abordarlos desde el punto de vista de 2ª persona, es una de las ideas centrales de Strawson, quien sostiene que cuando cambiamos de un punto de vista de 2ª persona de responsabilidad mutua, a una perspectiva de 3ª persona de observación, y, cuando vemos a los demás con miras a explicar y predecir su conducta en lugar de responder a su comportamiento, representa frecuentemente un movimiento de colocar a alguien fuera de la comunidad de agentes morales, ya sea temporalmente (ej. "él no es él mismo, recientemente perdió a su padre"; "aún no pueden evaluar completamente las consecuencias de sus acciones, sólo tienen catorce años") o permanentemente (ej. "su demencia hace que su comportamiento sea completamente inapropiado") (ver Strawson, 2008, p. 25).

En sentido concomitante, recordemos que, desde el enactivismo, la conexión afectiva surge de los involucramientos en segunda persona, lo cual es constituyente de la moral. Aunado a ello, como se vio, hay evidencia que sugiere que las interacciones en 2a persona forman una estructura autónoma que repercute normativamente en la agencia individual, así, la perspectiva de 2a persona en el encuentro con el otro puede modificar y determinar parte de la conducta de los participantes, con lo que arribamos al planteamiento de Colombetti y Torrance (2009) : lo que cada uno de nosotros hace en relación con el otro es parcialmente estructurado y caracterizado en términos interindividuales e interpersonales⁸².

En realidad, como ya he mencionado, considero que este aspecto de la moralidad no lo negaría la perspectiva del NMI, pero dados sus compromisos teóricos, lo explicaría desde una perspectiva computacionalista. Por ello, me interesa explicarlo y detallarlo a partir del enactivismo. Este movimiento al enactivismo me parece plausible por ser una teoría que explica y otorga autonomía al dominio de la interacción social entre agentes cognitivos. Al percibir la presencia del otro, en una perspectiva de 2a persona, se generan significados y conductas que se van acoplando momento a momento de la interacción, así, la interacción constituye un dominio cognitivo irreductible a la agencia individual, *i.e.*, emergen normas y significados que no están disponibles en un dominio exclusivamente individual.

Se trata, pues, de proponer una integración coherente que respete los principios de cada marco conceptual y sus modos de cognición, lo cual es posible al postular una arquitectura cognitiva del tipo PI, que si bien, aún no ha sido adoptada por la ciencia cognitiva, sí podría ser un camino prometedor de investigación, capaz de hacer justicia a los postulados de la ciencia cognitiva computacionalista y enactiva, respectivamente.

Por último, considero útil precisar cómo este modelo abordaría las dificultades presentadas.

4.10. Respuestas a las dificultades presentadas: articulando argumentos

4.10.1. Evolución

Respecto al argumento evolutivo, se vio que es posible articular plausiblemente los planteamientos de la escuela de Santa Bárbara y de la construcción de nicho si se recurre a los postulados de la genómica cultural; enfoque que intenta dar cuenta de la genuina influencia recíproca entre los genes y la cultura.

⁸² Considérese además que los juicios morales pueden variar dependiendo de la relación social que existe entre los agentes (Mammen, Köymen y Tomasello 2020).

Bajo este enfoque, se intentó ofrecer una descripción de *grano grueso* acerca de un tipo de conocimiento moral general dotado evolutivamente. En particular, se vio que dicho conocimiento pudo resultar de la interacción dinámica de dos herencias, la genética y la ecológica, lo que potenciaría un marco definicional que involucraría instintos paternos, filiales, de socialidad y empatía, integrados con la selección de mecanismos subyacentes a nuestras orientaciones normativas generales (núcleo biológico de rasgos heredables), que, en un ajuste con el nicho dinámico propiciado por las prácticas de aloparentalidad, la imitación, la intersubjetividad y la comunicación simbólica, entre otros, da cabida a pensar en la emergencia de las capacidades socio-morales humanas y con ello en el habilitamiento de nuestras nociones normativas amplias que surgieron de las estrategias exitosas o perniciosas de nuestros antepasados, lo cual constituiría nuestras clasificaciones compartidas acerca de lo que consideramos correcto-incorreto, aceptable-inaceptable, etc., mismas que en origen son objetivas pero multivalentes. Aquí, tómease en cuenta que el criterio de objetividad es crucial para evitar la apropiación cultural de los fundamentos morales, lo que ha sido blanco de críticas para las cuentas normativas relativistas.

4.10.2. Cognición

Puntualmente, se vio que, si aceptamos la arquitectura cognitiva PI, es posible contar con diversas estrategias para abordar diversas tareas, de tal manera que aquí se puede hablar plausiblemente tanto del uso de sustratos funcionalistas como del uso de sustratos corporizados.

Ello es posible debido a la naturaleza de PI que, por una parte, es compatible con la súper-modularidad computacionalista y es habilitada por representaciones, por procesamientos de información guiados por principios e involucra sustancialmente la información pragmática; y por otra parte, es compatible con los principios enactivos y es habilitada principalmente por la (inter)acción con el entorno efectuada a través de bucles de percepción-acción, lo que permite fundamentalmente el aprendizaje mediante la experiencia desde un enfoque relacional.

Así, pues, si mis consideraciones son correctas, pienso que la propuesta de la arquitectura cognitiva PI puede soportar plausiblemente los principios de cada marco conceptual, con sus respectivas metodologías, sus dos entidades teóricas núcleo (representación y acción) y por lo tanto sus dos modelos cognitivos (percepción-cognición-acción y percepción-acción), lo que da lugar a un crisol cognitivo más abarcador capaz de incluir diversas

estrategias cognitivas disponibles para ser empleadas al momento de enfrentarnos al inmenso y complejo mundo socio-moral que demanda la vida humana.

Por supuesto, esto apenas intenta ser una propuesta, pues aún faltaría mucho trabajo empírico y conceptual por hacer para habilitarla, no obstante, pienso que puede ser una alternativa genuina por explorar.

4.10.3. Cognición social

El modelo de la cognición social computacionalista-enactiva intenta ser una alternativa capaz de abordar la perspectiva de 2ª y de 3ª persona, según lo exija la situación, ya sea para atribuir estados mentales, o bien, para ser participantes involucrados en la interacción. Con ello, la pretensión es superar las dificultades presentadas por cada marco conceptual.

Por un lado, se hizo ver que la dificultad de la cognición social computacionalista se concentra en dos problemas: 1) debido a que la percepción es neutra, no explica cómo hacemos sentido de nosotros mismos y de los otros como *otros*, y; 2) concibe que el significado de la acción humana se encuentra en los estados mentales, los cuales son inobservables y por tanto inaccesibles a la percepción, con lo que se arriba al “problema de las otras mentes”.

De acuerdo con el modelo socio-cognitivo computacionalista-enactivo ambos problemas se disuelven. El problema (1) se disolvería debido a que un factor constitutivo de la percepción enactiva, es que el agente experimenta el comportamiento corporal de los demás como expresión de su alteridad y de su propia autonomía, de tal manera que en la interacción se percibe al otro como *otro* agente autónomo en su propio derecho.

En sentido relacionado, el problema (2) se disolvería debido a que en la parte enactiva, el significado del comportamiento no está localizado en estados o procesamientos que portan información dentro del sistema (como estados mentales ocultos), sino que el significado *surge* en la interacción misma, por tanto, en principio, las mentes de los otros y sus estados se nos presentan directamente y no son enteramente inobservables. Como se vio, en el proceso de la interacción, mediante los acoplamientos sensorio-motores y el involucramiento emocional, experimentamos inmediatamente al otro como *otro*, conocemos su mente y a su vez somos afectados y respondemos en consecuencia, por tanto, en las acciones de los otros percibimos sus significados.

Ahora bien, del lado enactivo, el problema surge por su visión exclusivamente interaccionista, pues para hablar de la cognición socio-moral en toda regla, se requiere conocer los procesamientos representacionales que la habilitan y dan lugar a ella, pues el aspecto perceptual y relacional del nivel supraindividual no es suficiente para explicarla. En particular, se trata de hablar de los procesamientos de información cargados de contenido que en su interacción dan lugar a la cognición socio-moral.

Para resolverlo, he recurrido a la arquitectura cognitiva súper-modular, la cual explica contundentemente los procesamientos funcionalistas-cognitivos del sujeto a partir de su naturaleza altamente interactiva que, a su vez, es capaz de involucrar de manera relevante la información pragmática.

Entonces, en mi propuesta, tanto los procesos de información representacionales y guiados por principios, como el aspecto perceptual y relacional del nivel supra-personal son parte constitutiva de la cognición socio-moral.

4.10.4. Cognición moral

En línea con lo anterior, se hizo ver (mediante la evidencia empírica presentada en los capítulos 2 y 3) que el espectro de la cognición moral es amplio y complejo, por lo que requerimos de la súper-modularidad computacionalista para dar cuenta satisfactoriamente de los procesos representacionales implicados en ella, pero también, he propuesto que se requiere de una metodología interaccionista para dar cuenta de la misma. Esto debido a que, en mi interpretación, es necesario otorgar autonomía a los procesos supra-personales desde su aspecto relacional, lo que considero que debe ganar un lugar en el estudio de la moral.

De ello, tendríamos un modelo integracionista de la cognición moral, el cual se beneficiaría de diversas estrategias, tales como la toma de perspectiva, el uso de ToM, el entendimiento intencional, el uso de representaciones, de heurísticas y de razonamiento basado en principios morales, etc., pero también, de estrategias corporizadas más básicas, las cuales favorecen la acción pre-reflexiva y los juicios morales en 2ª persona.

Concerniente a otorgar autonomía al nivel supra-personal y no sólo pensarlo como computado representacionalmente, implica estudiar la moral como estructura social desde una metodología relacional, lo cual, he argumentado, contribuye significativamente al menos a tres aspectos de la misma: 1) permite develar su aspecto situado mediante los

significados compartidos, 2) explica la co-constitución agente-agente (lo que da lugar a la dependencia recíproca), y, 3) posibilita explicar el papel de un agente *activo* genuinamente empotrado en un contexto permeado de normatividades y prácticas compartidas.

En concreto, el lado computacionalista, que es habilitado por la arquitectura súper-modular, explica contundentemente los procesamientos de información cargados de contenido que dan lugar a la cognición moral, los cuales son principalmente: ToM, entendimiento de intenciones, entendimiento de emociones, principios morales pre-lingüísticos y empatía. La interacción entre ellos es lo que habilita la competencia moral humana.

Y, el lado enactivo, que es habilitado por acoplamientos sensorio-motores guiados por la percepción y arraigado en la perspectiva de 2a persona (con lo que otorga a los agentes interactuantes un papel de participación y grados de responsabilidad moral), da lugar al menos a tres aspectos constituyentes de la cognición moral: 1) a un *tirón* moral fenomenológico sentido en el cuerpo vivo, 2) a la normatividad co-regulada que emerge de la interacción, y, 3) a la heteronormatividad del contexto socio-cultural.

De igual modo, he planteado que las contribuciones precisas del enactivismo son las siguientes: 1) explica cómo hacemos sentido del otro como *otro*, lo que no es trivial, sino que es una característica constituyente de la moral, 2) explica cómo nos auto-individuamos como agentes autónomos, 3) a partir de (1) y (2), se obtiene una perspectiva des-centrada del mundo que forma parte de nuestra perspectiva moral, y, 4) explica la intencionalidad compartida (mínima) y sus mecanismos subyacentes, lo que contribuye a formar una plataforma constituyente del desarrollo moral.

Así, pues, debido a que cada marco conceptual se enfoca en diferentes aspectos de la cognición socio-moral, es plausible pensar que ambos tienen pertinencia y que su complementación puede resultar en una explicación más abarcadora del fenómeno a estudiar: el computacionalismo explica los procesamientos representacionales que dan lugar a la cognición moral, y el enactivismo, a partir de su concepción de agencia e interacción, captura las dimensiones fenomenológicas, interactivas, corporizadas y afectivas de la misma. De esta manera, podríamos contar con un modelo que puede trabajar en diversos niveles, lo que considero que presenta ventajas frente a otras explicaciones de carácter excluyente, pues permitiría contar con diversas estrategias disponibles para efectuar las tareas relevantes.

4.10.5. Computacionalismo y enactivismo en tareas socio-morales

Si mis consideraciones han sido correctas, el computacionalismo y el enactivismo trabajarían en tareas socio-morales de la siguiente manera:

El enactivismo se encargaría, al menos, de la acción moral pre-reflexiva que surge del confrontamiento directo con el mundo, la cual es favorecida por bucles de percepción-acción y se enraíza en una perspectiva de 2ª persona. Asimismo, contribuiría a la ejecución de los juicios morales efectuados en 2ª persona à la Strawson.

Y, el computacionalismo, por su parte, se encargaría de todo lo demás, es decir, de los demás comportamientos y juicios morales, los cuales recurren esencialmente a representaciones, a razonamiento basado-en-principios, a la perspectiva de 3ª persona, al uso de ToM, al entendimiento de intenciones, al uso de heurísticas rápidas y frugales y, en general, a todo aquel templete que compone los procesamientos de información cargados de contenido que en interacción dan lugar a la cognición socio-moral.

4.11. Normatividad

Se vio que el computacionalismo brinda el argumento a favor de un conocimiento moral general dotado evolutivamente, lo que asegura la objetividad moral, y, el enactivismo, debido a que surge de la interacción y se instituye como práctica social, es capaz de explicar el aspecto *situado* de la normatividad moral.

Al respecto he argumentado que estos dos aspectos en conjunto permiten postular una herramienta evaluativa capaz de generar juicios situados pero anclados a criterios objetivos, lo que tal vez pueda invitar a nuevos acercamientos evaluativos respecto a dilemas morales.

En particular, concierne a la dificultad del lado computacionalista en la que he propuesto que requerimos también de una metodología interaccionista, he defendido que es plausible otorgar autonomía a los procesos supra-personales, movimiento que contribuye significativamente al menos a los tres aspectos de la normatividad moral que he mencionado⁸³.

⁸³ 1) permite develar su aspecto situado mediante los significados compartidos, 2) explica la co-constitución agente-agente (lo que da lugar a la dependencia recíproca), y, 3) posibilita explicar el papel de un agente activo genuinamente empotrado en un contexto permeado de normatividades y prácticas compartidas.

Concerniente al enactivismo, he sostenido que su dificultad refiere a la necesidad de recurrir al argumento del conocimiento moral general evolutivamente dotado, con lo que se asegura parsimoniosamente la objetividad moral y con ello se evita la apropiación cultural, lo que además es consistente con su principio moral básico enactivo del *cuidado de las autonomías*.

Como consideré anteriormente, desde un mapeo de *grano grueso*, se puede decir que la parte enactiva también emplearía el conocimiento moral general basado en principios, en tanto que es guiado fundamentalmente por el principio básico del *cuidado de las autonomías* (el cual surge de sus propios fundamentos), mismo que podría beneficiarse del compuesto de orientaciones normativas amplias y generales. Arguyo que esto es posible debido a que la arquitectura PI soporta que las estrategias trabajen cooperando o en entera independencia.

Para ser clara, si bien, el enactivismo no recurre principalmente a un razonamiento basado en principios, desde una postura moderada, se puede aceptar que el cuidado de las autonomías se suma plausiblemente a los principios morales universalmente compartidos. En principio, se podría conceder que este principio provendría (al menos en parte) también de aquella arquitectura heredada que nos instruye en algún grado con propensiones morales. Tal propensión parece beneficiarse de aquello que ha permanecido estable a lo largo del tiempo, y en torno a ello se ha argumentado que lo que se mantiene estable es lo necesario. En este sentido, en tanto que el “cuidado de las autonomías” podría atribuirse a aquello que benefició la *aptitud* humana, es plausible sumarlo a lo que ha permanecido estable en el tipo de arquitectura heredada y como tal formaría parte de nuestra dotación cognitiva. Ahora bien, los mecanismos y sustratos enactivos, aunque no coincidan con aquellos postulados por el computacionalismo, es posible que a partir de la matriz relacional expuesta en la Fig.2, el principio moral enactivo sea parte de lo que se mantiene como objetivo y general (marcado por la flecha), mientras que es atravesado y moldeado por las intervenciones inter-individuales y supra-personales (marcado por la espiral), lo que integraría la interacción del repertorio de esquemas sensoriomotores adquiridos mediante la historia de interacciones del agente, con aquellos rasgos que conforman nuestra arquitectura cognitiva dotada evolutivamente.

Así, pues, he considerado que la fusión de la normatividad computacionalista y enactiva posibilitaría una útil herramienta evaluativa capaz de abordar algunos dilemas morales.

En conclusión, si mis consideraciones han sido correctas, habría la posibilidad de un genuino diálogo entre el computacionalismo y el enactivismo que permita pensar en un marco integrativo capaz de abordar los principios y fundamentos de ambos, y con ello arribar a un modelo cognitivo no exclusivista y más abarcador. Pienso que la integración de enfoques puede ser un camino fructífero para comprender mejor la riqueza y la complejidad de la mente humana.

Por supuesto, la propuesta presentada requiere más trabajo empírico y conceptual, pues con la investigación realizada hasta aquí muchas cuestiones quedan abiertas, pero si mis interpretaciones son correctas y en efecto, se han sorteado adecuadamente las dificultades que presentan ambos marcos conceptuales, mi propuesta integrativa puede ser una alternativa genuina a la explicación del fenómeno socio-moral.

5. Conclusiones

En la presente investigación doctoral abordé dos principales teorías de la ciencia cognitiva, el computacionalismo y el enactivismo, respectivamente, con el objetivo de analizar sus postulados teóricos-empíricos para lograr arribar a una explicación integrativa de la cognición socio-moral humana.

En el capítulo 1, presenté la cartografía de los debates relevantes que han precedido a ambos marcos conceptuales, en donde abordé los cuatro temas centrales que guiaron la investigación: 1) evolución, 2) cognición, 3) cognición social y 4) cognición moral. Concerniente a la “Evolución”, expuse los acercamientos de las tres principales escuelas de la psicología evolutiva (la hipótesis de la mente modular, del cerebro social y del cerebro cultural), debido al gran impacto que han tenido en el progreso de los estudios de las ciencias cognitivas. En ello, se hizo ver que el computacionalismo es consistente con los principios de la mente modular, mientras que el enactivismo con la hipótesis del cerebro cultural, en específico, con los principios de la construcción de nicho, lo que deriva en hablar de dos herencias; la genética y la ecológica.

En la sección de “Cognición”, expuse los fundamentos de ambos marcos teóricos y los modelos cognitivos que habilitan a cada uno, en donde se vieron claramente las divergencias metodológicas y conceptuales que prevalecen entre ellos, los cuales son los siguientes: 1) en el computacionalismo hay una separación modal entre la percepción y la acción, mientras que en el enactivismo no la hay, pues la cognición es efectuada a partir de bucles continuos de percepción \leftrightarrow acción; 2) el computacionalismo parte de un individualismo metodológico (en tanto que su principal unidad de análisis es el sujeto individual), mientras que el enactivismo parte de una metodología relacional que suele defender una epistemología interaccionista; 3) el computacionalismo concibe la cognición como procesamientos de información regidos por reglas y los corporizados como procesos de hacer-sentido habilitados por las interacciones agente-entorno, y; 4) el computacionalismo trata con *representaciones* y contenidos semánticos para explicar la cognición, mientras que el enactivismo habla de *acción* y de un continuo entre la mente y la vida para dar cuenta de la misma.

Respecto a la “Cognición social”, presenté el panorama conceptual en el que se han situado los debates relevantes en torno a explicar las competencias socio-cognitivas humanas, en donde la disputa se traza de manera general entre dos concepciones antagónicas:

atribucionistas vs corporizados. Aquí, se hizo ver que, derivado de sus metodologías, el computacionalismo explica la cognición social mediante el proceso de atribución mental, mientras que los enactivos priman la interacción como el recurso explicativo de la misma, además de poner en primer plano la percepción directa del otro. De ello, los atribucionistas defienden una perspectiva epistémica de 3ª persona con lo que otorgan un papel de *observador* a los individuos, mientras que los corporizados defienden una postura de 2ª persona, con lo que otorgan un papel de *participante involucrado* a los interactuantes.

Concerniente a la “Cognición moral”, presenté los planteamientos de la psicología del desarrollo moral, en donde se vio que en los años 2000 surgió un giro metodológico que permitió nuevos acercamientos al estudio de la misma, lo que dio lugar al “nuevo paradigma de una moral innata”. Asimismo, expuse los acercamientos enactivos a su estudio y destacué lo *moderado* de sus pretensiones, pues el enactivismo busca *complementar* y no *sustituir* las explicaciones disponibles acerca de la cognición moral.

En el capítulo 2, abordé explícitamente la postura computacionalista y con ello expuse la arquitectura *súper-modular* en la que me baso para habilitar esta parte de mi propuesta. En cuanto a los cuatro temas eje que guiaron la investigación, en la sección de “Evolución” se planteó que la selección natural explica la presencia de instintos sociales (incluida la empatía como piedra angular), así como también, la selección de ciertos mecanismos que están detrás de nuestras orientaciones normativas generales, lo que, en conjunto, constituiría un núcleo biológico de rasgos heredables (herencia genética).

Concerniente a la “Cognición”, se presentó el procesamiento cognitivo-funcionalista, el cual, funciona de la siguiente manera: *input* (perceptual) → cognición → *output* (toma de decisiones que deriva en conductas manifiestas). No obstante, se hizo notar que, al adoptar la arquitectura *súper-modular*, los procesamientos ya no son lineales, sino, multidireccionales y altamente interactivos.

En la sección de la “Cognición social computacionalista”, expuse el proceso que le habilita en general y destacué los problemas que presenta su versión modularista-innatista, lo que permitió ver la necesidad de recurrir a una arquitectura del tipo *súper-modular* para lograr superarlos. En la sección de la “Cognición moral computacionalista”, se presentó la arquitectura *súper-modular* y a partir de ella se explicaron los procesamientos que le habilitan, los cuales son procesamientos de información cargados de contenido, que en su continua interacción dan lugar a la cognición moral (Teoría de la mente, principios morales

prelingüísticos, entendimiento intencional, entendimiento emocional y empatía, principalmente). Asimismo, se expuso el amplio cuerpo de trabajo empírico que lo sustenta como capacidad pre-lingüística (no aprendida), lo que sugiere que es parte de la dotación evolutiva de nuestro aparato cognitivo y, como tal, se plantea que asegura la objetividad moral.

Así, he planteado que las “implicaciones normativas computacionalistas” derivan en sugerir la presencia de un tipo de conocimiento moral general que nos orienta normativamente hacia nociones de lo que clasificamos como correcto-incorrecto, bueno-malo; nociones universalmente compartidas, lo que permite asegurar la objetividad moral por ser parte de nuestra dotación natural.

En el capítulo 3, presenté el enfoque enactivo, en el que busqué hacer ver lo valioso de adoptar una metodología *relacional* que otorga autonomía a los procesos supra-personales y se enfoca en el nivel perceptual. De ello, he planteado que sus principales contribuciones son las siguientes: 1) concerniente a la “Evolución”, se vio que la herencia ecológica se constituyó a partir de un nicho dinámico articulado y heredado por el aprendizaje cultural (o social), el cual definió el escenario que dio lugar a la emergencia de las capacidades socio-morales humanas; 2) respecto a la “Cognición”, expuse que ésta es entendida como procesos de *hacer-sentido* habilitados por las interacciones agente-entorno, de lo que surgen significados y se instauran normatividades, así como también, presenté su modelo cognitivo que se efectúa a manera de bucles continuos de percepción \leftrightarrow acción, donde no hay separación modal entre ellas; 3) en la “Cognición social enactiva”, mostré las condiciones de la “interacción socio-cognitiva” que dan lugar al surgimiento de la normatividad social, la cual es co-regulada y situada; 4) en la sección de la “Normatividad moral enactiva” expuse su principio moral básico concerniente al *cuidado de las autonomías* de los agentes, mismo que posee un pluralismo de realizabilidad que permite efectuar juicios morales *situados*, y; 5) presenté la evidencia empírica que apuntala las intervenciones corporizadas en la cognición socio-moral, las cuales contribuyen principalmente a la acción moral pre-reflexiva y a los juicios morales en 2ª persona.

Por último, en el capítulo 4 presenté la articulación de los argumentos expuestos y describí la arquitectura cognitiva PI que propongo, con lo que se habilita una explicación integrativa de la cognición socio-moral, que involucra ambos modos de cognición (computacionalista y enactivo), respetando sus principios y metodologías; lo que considero una genuina alternativa explicativa respecto al fenómeno a estudiar.

6. Bibliografía

- Aizawa, K. (2014). The Enactivist Revolution. *AVANT*, Vol. V, No. 2.
- Álvarez, J. R. (2013). Selección natural y construcción de nicho: una ¿dialéctica? *Evolucionista. Contrastes, Revista Internacional de Filosofía, Suplemento*, 18, pp. 343-355.
- Arbib, M. (2005). From Monkey-like Action Recognition to Human Language: An Evolutionary Framework for Neurolinguistics. *Behavioral and Brain Sciences* 28, pp. 105 – 168.
- Asch, S. (1952). *Social psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Barrett, N. (2015). The Normative Turn in Enactive Theory: An Examination of Its Roots and Implications. *Springer Science+Business*.
- Barkow, J., Cosmides, L. y Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. NY: Oxford University Press
- Baron-Cohen, S. (1997). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge MA: MIT Press.
- Bechtel, W. (2008). *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York, NY: Routledge.
- Bigelow, A. y Rochat, P. (2006). Two-month-old infants' sensitivity to social contingency in mother-infant and stranger-infant interaction. *Infancy* 9(3), pp.313-25.
- Boden, M. (2006). *Mind as machine*. Oxford: Oxford University Press
- Bohl, V. y van den Bos, W. (2012). Toward an integrative account of social cognition: marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience*, 6, pp. 1–15.
- Bohl, V. (2014). *How do we understand others? Beyond theories of mindreading and interactionism* (Doctor in Philosophy): University of Tartu Press
- Bramer, M. (2010). The Importance of Personal Relationships in Kantian Moral Theory: A Reply to Care Ethics. *Hypatia*, Volume 25, Issue 1: Special Issue: Current Work in Feminist Ethics and Social Theory, pp. 121 - 139.
- Bråten, S. (2007). *On Being Moved: From Mirror Neurons to Empathy*. Amsterdam: John Benjamins.
- Brown, J. E. (2012). "The Evolution of Symbolic Communication: An Embodied Perspective." PhD Thesis, University of Edinburgh.
- Brown, J. E. (2012). *The Evolution of Symbolic Communication: An Embodied Perspective*. PhD Thesis, University of Edinburgh.
- Brüne, M. (2007). On human self-domestication, psychiatry, and eugenics. *Philosophy, Ethics, and Humanities in Medicine*, 2(21). doi:10.1186/1747-5341-2-21
- Brunsteins, P. (2010). *La psicología folk: teorías, prácticas y perspectivas*. Ediciones del signo, Argentina.
- Burkart, J.M., Hrdy, S.B. y Van Schaik C.P. (2009). Cooperative breeding and human cognitive evolution. *Evol Anthropol*, 18, pp. 175-186.
- Carey, S. y Spelke, E. (1994). Domain specific knowledge and conceptual change. En L. A. Hirschfeld & S. A. Gelman. (Eds.), *Mapping the mind. Domain specificity in cognition and culture* (pp. 243-284). Cambridge: Cambridge University Press. [Trad. cast. en 2 volúmenes de A. Ruíz: Conocimiento dominio específico y cambio conceptual. En *Cartografía de la mente. La especificidad de dominio en la cognición y la cultura*. Barcelona: Gedisa, 2002]
- Carpendale, J., Frayn, M. y Kucharczyk, P. (2017). The social formation of human minds. En J. Kiverstein, (Ed.), *The Routledge Handbook of Philosophy of the Social Mind*. NY: Routledge.
- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32, pp. 121–182.
- Chen, C. y Moyzis, R. (2018). Cultural Genomics: Promises and Challenges. *Journal of Cross-Cultural Psychology*, 49(5), pp. 764–788.
- Churchland, P. (1992). *Materia y Conciencia: Introducción contemporánea a la filosofía de la mente*. Gedisa (Filosofía).
- Cohn, J. F. y Tronick, E. Z. (1989). Specificity of infants' response to mothers' affective behavior. *Journal of the American Academy of Child and Adolescent Psychiatry* 28, pp.242-8.
- Colombetti, G. y Torrance, S. (2009). Emotion and Ethics: An Inter-(en)active Approach. *Phenomenology and the Cognitive Sciences*, 8(4), pp. 505-526.
- Colombo, M. (2012). Constitutive relevance and the personal/subpersonal distinction. *Philos. Psychol*, pp. 1-24.

- Copeland, J. y Proudfoot, D. (1996). On Alan Turing's anticipation of connectionism. *Synthese*, 108, pp. 361-377.
- Costall, A., Leudar, I. & Reddy, V. (2006). Failing to see the irony in mind-reading. *Theory & Psychology*, 16(2), pp.163-7.
- Csibra, G. y Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), pp.148-153.
- Currie, G. y Sterelny, K. (2000). How to think about the modularity of mind-reading. *The Philosophical Quarterly*, 50(199), pp. 145–160.
- Davidov, M., Hanania, R. y Zahan-Waxler, C. (2011). Empathy development from 8 to 16 months: Early signs of concern for others. *Infant Behavior and Development*, Volume 34, Issue 3.
- Davies, M. y Stone, T., (Eds.). (1995). *Mental simulation: evaluations and applications*. Oxford/Cambridge, MA: Blackwell.
- De Bruin, L. y Kästner, L. (2012). Dynamic Embodied Cognition. *Phenomenology and the Cognitive Sciences*, 11, pp. 541–563.
- Dennett, D. (1969). *Content and Consciousness*. London: Routledge.
- De Jaegher, H. y Di Paolo, E. A. (2007). Participatory Sense-Making: An Enactive Approach to Social Cognition. *Phenomenology and the Cognitive Sciences*, 6(4).
- De Souza, N. (2013). Pre-reflective ethical know-how. *Ethic Theory Moral Prac.* 16, pp. 279–294.
- Di Paolo, E., Rohde, M., De Jaegher, H. (2010). Horizons for the Enactive Mind: Values, Social Interaction, and Play. En: Stewart, J., Gapenne, O. y Di Paolo, E. A. (Eds.), *Enaction: To-towards a New Paradigm for Cognitive Science*, Cambridge, MA: MIT Press, pp. 33-87.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and The Cognitive Sciences*, 4, pp. 429–452.
- Di Paolo, E., Buhrmann, T. y Barandiaran, X. (2017). *Sensorimotor Life: An Enactive Proposal*. Oxford University Press, EUA.
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Human Culture*. Cambridge, MA: Harvard University Press.
- Doris, John, Stich, S., Phillips, J. y Walmsley, L. (2020). Moral Psychology: Empirical Approaches. *The Stanford Encyclopedia of Philosophy (Spring 2020 Edition)*, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2020/entries/moral-psych-emp/>>.
- Edwards, B. (2017). *The Challenge of Reconciling Analytic and Continental Moral Philosophy*. Artículo preparado para su presentación en University College Dubli. Recuperado en: <https://medium.com/@benjaminjedwards/the-challenge-of-reconciling-analytic-and-continental-moral-philosophy-4deb2421e10b>
- Eisenberg, N., Spinrad, T. y Knafo-Noam, A. (2015). Prosocial development. En W. Damon, W. (Ed.) *Handbook of Child Psychology and Developmental Science (7th edition)*, Hoboken, NJ: Wiley.
- Enesco, I. y Delval, J. (2006). Módulos, dominios y otros artefactos. *Infancia y Aprendizaje*, 29(3), pp. 249-267.
- Engels, E.M. (2016). The Roots of Human Morals and Culture in Pre-Human Sympathy: Charles Darwin's Natural and Cultural History of Morals. En Etzelmüller, G. y Tewes, C., (Eds.). (2016). *Embodiment in Evolution and Culture*. Alemania: Mohr Siebeck.
- Etzelmüller, G. y Tewes, C., (Eds.). (2016). *Embodiment in Evolution and Culture*. Alemania: Mohr Siebeck.
- Farroni, T., Csibra, G., Johnson, M. y Simion, F. (2002) Eye contact detection at birth. *Proceedings of the National Academy of Sciences* 99(14), pp.9602-5.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47, pp. 164–203.
- Ferreira, F. y Patson, N.D. (2007). The 'Good Enough' Approach to Language Comprehension. *Language and Linguistics Compass*, 1/1–2, pp. 71–83.
- Flynn E., Laland, K., Rachel L., Kendal y Jeremy R. (2013). *Target Article with Commentaries, Developmental niche construction*, Blackwell Publishing.
- Fodor, J. (1991). Fodor's Guide to Mental Representation. En Greenwood, J., (Ed.), *The Future of Folk Psychology*, Cambridge University Press.
- Fodor, J. (1978). Las actitudes proposicionales. En Rabossi, E., (Ed), 1995, *Filosofía de la Mente y Ciencia Cognitiva*, Ediciones Paidós. pp. 173-203.
- Fodor, J. (1983). *The Modularity of Mind*. MIT Press, Cambridge, MA.
- Fodor, J. (1968). The appeal to tacit knowledge in psychological explanation. *Journal of Philosophy*, 65, pp. 627-640.

- Fogel, A. (1993). *Developing through Relationships: Origins of Communication, Self and Culture*. New York: Harvester Wheatsheaf.
- Foot, P. (2002). Nietzsche: The Revaluation of Values. En *Virtues and Vices: and other essays in moral philosophy*, ed. Foot, P. UK: Oxford University Press, pp.81–95.
- Frijda, N. H. (1986). *The Emotions*. Cambridge University Press.
- Froese, T. y Di Paolo, E. (2009). Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences*, pp. 1-25.
- Froese, T. y Di Paolo, E. A. (2011). The Enactive Approach. *Theoretical Sketches from Cell to Society. Pragmatics & Cognition*, 19(1), pp. 1-36.
- Froese, T. (2012). From adaptive behavior to human cognition: a review of Enaction. *Adaptive Behavior*, 20(3), pp. 209–221.
- Froese, T. y Leavens, D. A. (2014). The direct perception hypothesis: Perceiving the intention of another's action hinders its precise imitation. *Frontiers in Psychology*, 5(65).
- Froese, T., Iizuka, H., Ikegami, T. (2013). From synthetic modeling of social interaction to dynamic theories of brain-body-environment-body-brain systems. *Behavioral and Brain Sciences*, 36(4), pp. 420-21.
- Froese, T. (2015). Enactive neuroscience, the direct perception hypothesis, and the socially extended mind. *Behavioral and Brain Sciences*.
- Fuchs, T. y De Jaeger, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8(4), pp. 465–486.
- Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies*, 8(5–7), pp. 83–108.
- Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, & Psychology*, 11(3), pp. 199–217.
- Gallagher, S. (2005). Phenomenological contributions to a theory of social cognition. *Husserl Studies*, 21(2), pp. 95–110.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford University Press, USA.
- Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2:3, pp. 353 - 365.
- Gallagher, S. (2008). Direct perception in the intersubjective context. *Conscious. Cogn.* 17, pp. 535–543.
- Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press, UK.
- Gallese, V. (2009). Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanalytic Dialogues*, 19(5), pp. 519–536
- Gamble, C., Gowlett, J. y Dunbar, R. (2014). *Thinking big: How the evolution of social life shaped the human mind*. NY: Thames & Hudson.
- Gangopadhyay, N. y Schilbach, L. (2011). Seeing minds: a neurophilosophical investigation of the role of perception-action coupling in social perception. *Soc. Neurosci.* 7, pp. 1-14.
- Garcia-Ramirez, E. (2019). *Open Compositionality: Toward a New Methodology of Language*. MD: Lexington Books.
- Garcia-Ramirez, E. (2019). Morality, Cognition and the Language Analogy. En Garcia-Ramirez, E., *Open Compositionality*, MD: Lexington Books.
- Geraci, A. y Surian, L. (2011). The Developmental Roots of Fairness: Infants' Reactions to Equal and Unequal Distributions of Resources. *Developmental Science*, 14. pp. 1012-20.
- Gigerenzer, G., Todd, P. M. y the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford University Press, USA.
- Goldstein, M. H. y Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science* 19(5), pp. 515-23.
- Goldstein, M. H. y Schwade, J. (2010). From birds to words: Perception of structure in social interactions guides vocal development and language learning. En Blumberg, M. S., Freeman, J. H. y Robinson, S. R. (Eds.), *The Oxford Handbook of Developmental Behavioral Neuroscience*. Oxford University Press (pp. 708-29).
- Gopnik, A., y Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, Vol. 7, Numbers 1 and 2 Spring/Summer.
- Goswami, U. (2004). *Childhood Cognitive Development*, Blackwell Publishers.
- Grandy, R. (1973). Reference, Meaning and Belief, en *The Journal of Philosophy*, 70(14), pp. 439-452.

- Gray, K., y Schein, C. (2018). The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm. *Personality and Social Psychology Review*, Vol. 22(1) pp. 32-70.
- Gutiérrez-Martínez, F. (2005). *Teorías del desarrollo cognitivo*. McGraw-Hill, España.
- Hamlin, J. K., Wynn, K. y Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, pp. 450, 557.
- Hamlin, J.K., Wynn, K., Bloom, P. y Mahajan, N. (2011). How infants and toddlers reach to antisocial others. *Proceedings of the National Academy of Sciences*, 108, pp. 19931-19936.
- Hamlin, J. K. (2012). A Developmental Perspective on the Moral Dyad. *Psychological Inquiry*, 23, pp. 166–171.
- Hamlin, J.K. (2013). Moral judgment and action in preverbal infants and toddlers: Evidence for an innate moral core. *Current Directions in Psychological Science*, 22, pp.186-193.
- Hamlin, J.K., Ullman, T., Tenenbaum, J., Goodman, N. y Baker, C. (2013). The mentalistic basis of core social cognition: experiments in preverbal infants and a computational model. *Dev Sci*, 16(2), pp. 209-226.
- Hamlin, J. K. (2015). The infantile origins of our moral brains. En Decety, J. y Wheatley, T. (Eds.), *The moral brain: A multidisciplinary perspective* (pp. 105–122). Boston Review.
- Haugeland, J. (1990). The intentionality all-stars. *Philosophical Perspectives*, 4, pp. 383–427.
- Henderson, A. M. E. y Woodward, A. L. (2011). Let's work together: What do infants understand about collaborative goals? *Cognition*, 121(1), pp. 12–21.
- Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. NJ: Princeton University Press.
- Hernández-Chávez, P. y García-Campos, J. (2020). Modularizando y desmodularizando la cognición social: el caso de la detección de trampa. *Contrastes. Revista Internacional de Filosofía*, vol. XXV N°2, pp. 57-78.
- Heyes, C. (2019). Is morality a gadget? Nature, Nurture and Culture in Moral Development. *Synthese on The Cultural Evolution of Human Social Cognition*.
- Hobson, P. (2002). *The Cradle of Thought*. Palgrave-McMillan: Oxford.
- Horst, S. (2016). *Cognitive Pluralism*. The MIT Press: Cambridge, Massachusetts.
- Hrdy, S. B. (2009). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Cambridge, MA: Harvard University Press.
- Hutto, D. (2009). ToM rules, but it is not OK! En Leudar, I. y Costall, A. (Eds.), *Against Theory of Mind* (pp. 221–238). Basingstoke: Palgrave Macmillan.
- Jayawickreme, E. y Chemero, A. (2008). Ecological Moral Realism: An Alternative Theoretical Framework for Studying Moral Psychology. *Review of General Psychology*, Vol. 12, No. 2, pp. 118–126.
- Jones, S. (2000). *Darwin's ghost: The origin of species updated*. New York, NY: Ballantine Books.
- Joyce, R. (2006). *The Evolution of Morality*, MIT Press, UK.
- Joyce, R. (2018). *The Routledge Handbook of Evolution and Philosophy*. Routledge, NY.
- Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge: MIT Press.
- Keltner, D. y Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion* 13(5):505-21.
- Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.
- Kirchmair, L. (2017) Morality Between Nativism and Behaviorism. *The Journal of Theoretical and Philosophical Psychology*, Vol. 37, No. 4, pp. 230–260.
- Kittay, E. (2009). The personal is philosophical is political. A philosopher and mother of a cognitively disabled person sends notes from the battle ground. *Metaphilosophy*, 40(3–4), pp. 606–627.
- Kiverstein, J., (Ed.), (2017). *The Routledge Handbook of Philosophy of the Social Mind*. NY: Routledge.
- Knafo-Noam, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Phil. Trans. R. Soc.*
- Knudsen, B. y Liszowski, U. (2013). One-year-olds warn others about negative action outcomes. *Journal of Cognition and Development*, 14(3), 424–436.
- Kochanska, G. (2002). Mutually responsive orientation between mothers and young children: A context for early development of conscience. *Current Directions in Psychological Science*, 11, pp. 191-195.
- Kochanska, G., Koenig, J.L., Barry, R.A., Kim, S, y Yoon, J.E. (2010). Children's conscience during toddler and preschool years, moral self, and a competent adaptive developmental trajectory. *Developmental Psychology*, 46, pp. 1320-1332.
- Korsgaard, C. M. (2010). Reflections on the Evolution of Morality. Recuperado en: <http://dash.harvard.edu/handle/1/5141952>

- Kouba, L.J. y Muasher, J. (1985). Female circumcision in Africa: an overview. *African studies review: the journal of the African Studies Association*, Vol. 28, 1, pp. 95-110.
- Kovacs, A.M., Teglás, E. y Endress, A.D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330, pp. 1830-1834.
- Kriegel, U. (2007). Moral phenomenology: Foundational issues. *Phenom Cogn Sci*.
- Kugiumutzakis, G. (1998). Neonatal imitation in the intersubjective companion space. En Braten, S. (ed). *Intersubjective Communication and Emotion in Early Ontogeny*. Cambridge: Cambridge University Press. (pp. 63-88).
- Kuhlmeier, V., Wynn, K. y Bloom, P. (2003). Attribution of Dispositional States by 12-Month-Olds. *Psychological science*, 14, pp. 402-8.
- Lagattuta, K.H. (2005). When you shouldn't do what you want to do: Young children's understanding of desires, rules, and emotions. *Child Development*, 76, pp. 713-733.
- Laland, K. (2017). *Darwin's unfinished symphony: How culture made the human mind*. NJ: Princeton University Press.
- Lane, J.D., Wellman, H.M., Olson, S.L., LaBounty, J. y Kerr, D.C.R. (2010). Theory of mind and emotion understanding predict moral development in early childhood. *British Journal of Developmental Psychology*, 28, pp. 871-889.
- Lapsley, D.K. y Carlo, G. (2014). Moral development at the crossroads: New trends and possible futures. *Developmental Psychology*, 50(1), pp. 1-7.
- Leslie, A. M. (1994). ToMM, ToBY, and Agency: Core architecture and domain specificity. En Hirschfeld, L. A y Gelman, S. A. (Eds.), *Mapping the Mind*. Cambridge: Cambridge University Press, pp. 119-148.
- Leudar, I. y Costall, A. (2008). *Against theory of mind*. London: Macmillan.
- Liszkowski, U., Carpenter, M. y Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108(3), pp. 732-9.
- Loidolt, S. (2018). Experience and Normativity: The Phenomenological Approach. En Antonio Cimino y Cees Leijenhorst (Eds): *Phenomenology and Experience: New Perspectives*.
- Macintyre, A. (2013) [1981]. *After Virtue*. UK: Bloomsbury.
- Maiese, M. (2011). *Embodiment, Emotion, and Cognition*. Palgrave Macmillan.
- Mammen, M., Köymen, B. y Tomasello, M. (2020). Young children's moral judgments depend on the social relationship between agents. *Cognitive Development*.
- Martin, G. B. y Clark, R. D. (1982). Distress crying in neonates: Species and peer specificity. *Developmental Psychology*, 18(1), pp. 3-9.
- Meltzoff, A. y Moore, K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198(4312), pp.75-78.
- Métais, F. y Villalobos, M. (2021). Embodied ethics: Levinas' gift for enactivism. *Phenomenology and the Cognitive Science*, 20(3), pp.1-22.
- Michael, J. (2011). Interactionism and mindreading. *Review of Philosophy and Psychology*, 2(3), pp. 559-578.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences*, 11, pp. 143-152.
- Mikhail, J. (2011). *Elements of Moral Cognition: Rawls's Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*, New York: Cambridge University Press
- Mithen, S. (1998). *Arqueología de la mente: Orígenes del arte, de la religión y de la ciencia*. Crítica: Barcelona.
- Mojica, L. y Froese, T. (2019). On the spatiotemporal extensiveness of sense-making: Ultrafast cognition and the historicity of normativity. *Synthese* 198 (Suppl 1), pp. 447-460.
- Monsó, S. (2017). Morality Without Mindreading. *Mind & Language*, 32(3), pp. 338-357.
- Nagy, E. y Molnar, P. (2004) Homo imitans or homo provocans? Human imprinting model of neonatal imitation. *Infant Behavior and Development* 27(1), pp. 54-63.
- Nam, Y. E. (2018). The power structure in the perpetuation of female genital cutting in Kenya. *Asian Journal of Women's Studies*, 24(1), 128-139.
- Narvaez, D. (2014). *Neurobiology and the development of human morality: Evolution, culture, and wisdom*. W. W. Norton & Company.
- Ngarũiya Njambi, W. (2011). Irua Ria Atumia and Anticolonial Struggles among the Gĩkũyũ of Kenya: A Counternarrative on "Female Genital Mutilation." En Oyèwùmí, O. (Ed.), *Gender Epistemologies in*

- Africa Gendering Traditions, Spaces, Social Institutions, and Identities*. Palgrave Macmillan. (pp. 179–197).
- Nichols, S. y Stich, S. (1995). Folk Psychology: Simulation or Tacit Theory. En Davies y Stone (eds.), (1995). *Mental simulation: evaluations and applications*, Oxford / Cambridge, MA: Blackwell, pp. 123-158.
- Odling-Smee, F.J., Laland, K., Feldman, M.W. (2003). *Niche Construction. The neglected process in evolution*. Princeton: Princeton University Press.
- Over, H. y Carpenter, M. (2009). Eighteen-Month-Old Infants Show Increased Helping Following Priming with Affiliation. *Psychological Science*: 20(10), pp. 1189-1193.
- Overton, W. F. (1998). Developmental Psychology: Philosophy, concepts, and methodology. En Damon, W. (Ed. Principal), Lerner, R. M. (Ed. Vol.), *Handbook of child Psychology. Theoretical models of human development*. Nueva York: John Wiley & Sons. (pp. 107-189).
- Pacherie, E. (2013). Intentional joint agency: shared intention lite. *Synthese* 190, pp. 1817-1839.
- Pater, J. (2019). Generative linguistics and neural networks at 60: Foundation, friction, and fusion. *Language* 95, pp. 41–74.
- Perconti, P. y Plebe, A. (2020). Deep Learning and Cognitive Science. *Cognition*, 203.
- Pescador, C. y Mojica, L. (2022). Making us autonomous: The enactive normativity of morality. *Topoi*, 41, pp. 257–274.
- Pescador, C. y Melchor, P. (2024). Nuevos paradigmas en filosofía moral: aproximaciones enactivas. En Martínez-Bohórquez, M., *Perspectivas contemporáneas de la teoría moral*, UAM, (en revisión).
- Premack, D. y Premack, A. J. (1997). Infants attribute value± to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, 9(6), pp. 848–856.
- Premack, D. (2004). Is Language the Key to Human Intelligence? *Perspectives*, Vol. 303.
- Quiroga-Méndez, M. (2013). El innatismo moral, un nuevo paradigma de desarrollo moral, aportaciones desde la cognición y la neurociencia. *Acción Psicológica*, 10(2), pp. 179-188.
- Ramenzoni, V. y Liszkowski, U. (2016). The Social Reach: 8-Month-Olds Reach for Unobtainable Objects in the Presence of Another Person. *Psychol Sci.*, 27(9), pp. 1278-85.
- Ramos-Arenas, J. (1992). Reflexiones sobre el innatismo. *Ideas y valores*, No.89.
- Raschka, S. y Mirjalili, V. (2017). *Python Machine Learning* (second edition). Packt Publishing
- Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology, theory of mind and simulation*. Palgrave Macmillan.
- Ravenscroft, I. (2004). Folk Psychology as a Theory. *Stanford Enciclopedia of Philosophy*. En: <http://plato.stanford.edu/entries/folkpsych-theory/>
- Reddy, V. (2000). Coyness in early infancy. *Developmental Science* 3(2):186-92.
- Reddy, V. y Morris, P. (2004). Participants Don't Need Theories Knowing Minds in Engagement. *Theory and Psychology* 14(5), pp. 647-665.
- Reddy, V. (2008). *How Infants Know Minds* (1 edition). Cambridge, Mass: Harvard University Press.
- Reddy, V. (2015). Joining intentions in infancy. *Journal of Consciousness Studies*. 22, 1-2, pp. 24-44.
- Rheingold, H. L. (1982). Little children's participation in the work of adults, a nascent prosocial behavior. *Child Development*, 53(1), pp. 114–125.
- Richerson, P. J., y Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: The University of Chicago Press.
- Rosas, A. (2011). La evolución de la moral contractual. *Ideas y valores*, (LX) 147, pp. 209-222.
- Ryle, G. (1946). Knowing How and Knowing That: The Presidential Address. *Proceedings of the Aristotelian Society* 46, pp.1 - 16.
- Sanso-Guerrero, J.A. (2018). Autoconciencia y desarrollo moral: Una visión general desde la teoría de la construcción del nicho (NCT), *Universitat de Girona*.
- Satne, G. y Salice, A. (2015). Helping Behaviour and Joint Action in Young Children. *Phenomenology and Mind*, 9, pp. 98-106.
- Satne, G. (2016). Joint and Various (collective) forms of intentionality. *Journal of Social Ontology*, 2(1), pp.105-116.
- Satne, G. y Salice, A. (2018). Shared Intentionality and the Cooperative Evolutionary Hypothesis. *Springer*
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R. y Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia* 44(5), pp.718-30.
- Schilbach, L., Eickhoff, S. B., Mojzisch, A. y Vogeley, K. (2008). What's in a smile? Neural correlates of facial embodiment during social interaction. *Social Neuroscience* 3(1), pp.37-50.
- Schilbach, L. (2010) A second-person approach to other minds. *Nature Reviews Neuroscience* 11(6):449.

- Schilbach, L., Timmermans, B., Reddy, V (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences* 36(4), pp.393-414.
- Schmidt, M. y Sommerville, J. (2011). Fairness Expectations and Altruistic Sharing in 15-Month-Old Human Infants. *PLoS one*, 6(10).
- Scholl, B. J. y Leslie, A. M. (1999). Modularity, development, and Theory of Mind. *Mind & Language*, 14(1), pp. 131–153.
- Segal, G. (1996). The modularity of theory of mind. En P. Carruthers y P. K. Smith (Eds.), *Theories of Theories of Mind* (pp. 141–157), Cambridge: Cambridge University Press.
- Segovia-Cuéllar, A. (2017). Reflexiones sobre la explicación evolutiva en ciencias cognitivas: El origen de la cognición social humana como estudio de caso. *Revista argentina de antropología biológica*, Volumen 19, Número 1.
- Singh, P. (2018). Contrasting Embodied Cognition with Standard Cognitive Science: A Perspective on Mental Representation. *Journal of Indian Council of Philosophical Research*, 36, pp. 125–149.
- Sober, E., y Wilson, D. S. (1999). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stein, E. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*, Clarendon Press.
- Steiner, P., Sewart, J. (2009). From Autonomy to Heteronomy (and Back): The Enaction of Social Life. *Phenomenology and the Cognitive Sciences*, 8(4), pp. 527-550.
- Sterelny, K. (2005). Made by Each Other: Organisms and their Environment. *Biology and Philosophy*, 20, pp. 21-36.
- Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. MIT Press.
- Sterelny, K., Joyce, R., Calcott, B. y Fraser, B. (Eds.) (2013). *Cooperation and its Evolution*. MIT Press.
- Stern, D. (1985). *The Interpersonal World of the Infant*. New York: Basic Books.
- Strawson, P.F. (1974). *Freedom and resentment, and other essays*, vol 595. Egmont Books, London.
- Strawson, P.F. (2008). Freedom and Resentment. En McKenna, M. y Russell, P. (Eds.), *Free Will and Reactive Attitudes*. Ashgate Publishing Company.
- Sutcliffe, A., Dunbar, R., Binder, J. y Arrow, H. (2012). Relationships and the Social Brain: Integrating Psychological and Evolutionary Perspectives. *British Journal of Psychology* 103 (2), pp. 149 – 168.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard UP: Cambridge, Mass.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav Brain Sci*, 28(5), pp. 675-91.
- Tomasello, M. y Carpenter, M. (2005). Imitation reading and imitative learning. En Hurley, S. y Chater, N. (Eds.), *Perspectives on imitation: From neuroscience to social science. Imitation, human development, and culture* (Vol. 2). Cambridge, MA: MIT Press.
- Tomasello, M. (2008). *Origins of Human Communication*. MIT Press: Cambridge, Mass.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Cambridge, MA: Harvard, Belknap.
- Todd, P.M., y Gazzinger, G. (2000). Précis of Simple heuristics that make us smart, *Behavioral and Brain Sciences*, 23, pp. 727–780.
- Torrance, S. y Froese, T. (2011). An Inter-enactive Approach to Agency: Participatory Sense-making, Dynamics, and Sociality. *Humana Mente*, 15, pp. 21-53.
- Trevarthen, C. (1977). *Descriptive analyses of infant communication behavior*. En Studies in Mother-Infant Interaction: The Loch Lomond Symposium, ed. H.R. Schaffer. London: Academic Press. (pp. 227-70).
- Trevarthen, Colwyn y Bullowa, M. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. *Before Speech* (Cambridge), pp. 321-347.
- Trevarthen, C. (2011). What Is It Like To Be A Person Who Knows Nothing? Defining the Active Intersubjective Mind of a Newborn Human Being. *Infant and Child Development*, 20 (1), pp. 119-135.
- Urban, P. (2014). Toward an expansion of an enactive ethics with the help of care ethics. *Front Psychol* 5, pp.1354.
- Urban, P. (2015). Enactivism and care ethics: merging perspectives. *Filozofia* 70(2), pp.119–129.
- van Grunsven, J. (2015). *Bringing Life in View: An Enactive Approach to Moral Perception*. (Doctor in Philosophy): The New School University.
- van Grunsven, J. (2017). Enactivism, Second-Person Engagement and Personal Responsibility. *Phenomenol Cogn Sci*, 17(1), pp. 131–156.

- Varela, F.J., Thompson, E. y Rosh, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT Press.
- Varela, F.J. (1991) Organism: a meshwork of selfless selves. En Boston Studies in the Philosophy of Science. *Springer*, pp. 79–107.
- Varela F.J. (1999). *Ethical know-how: action, wisdom, and cognition*. Stanford University Press, Redwood City.
- Waldenfels, B. (2016). *Grundmotive einer Phänomenologie des Fremden* (5th ed). Frankfurt a. M.: Suhrkamp.
- Wallot, S. y Van Orden, G. (2012). Ultrafast Cognition. *Journal of Consciousness Studies*, 19(5-6), pp. 141-160.
- Warneken, F., Chen, F. y Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77(3), pp. 640–663.
- Warneken, F. y Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11(3), pp. 271-294.
- Weber A, Varela F. (2002). Life after Kant: natural purposes and the autopoietic foundations of biological individuality. *Phenomenol Cogn Sci*, 1(2), pp. 97–125. [https:// doi. org/ 10.1023/A: 1020368120174](https://doi.org/10.1023/A:1020368120174)
- Wellman, H.M. y Miller, J.G. (2008). Including deontic reasoning as fundamental to theory of mind. *Human Development*, 51, pp. 105-135.
- Williams, B. (2000). Naturalism and Genealogy, En *Morality, Reflection and Ideology*, ed. Harcourt, E. UK: Oxford University Press, pp.148–161.
- Wilson, C. (2018). Evolution and Ethics: An Overview. En Joyce, R. (Ed.), (2018). *The Routledge Handbook of Evolution and Philosophy*. Routledge: NY and London.
- Wilson, R.A. y Foglia, L. (2017). Embodied Cognition. *The Stanford Encyclopedia of Philosophy (Spring 2017 Edition)*, Edward N. Zalta (ed.) <<https://plato.stanford.edu/archives/spr2017/entries/embodied-cognition/>>.
- Wimmer, H. y Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children´s understanding of deception, *Cognition*, 13, pp. 103-128.
- Wulf, C. (2016). The Creation of Body Knowledge in Mimetic Processes. En Etzelmüller, G. y Tewes, C., (Eds.). (2016). *Embodiment in Evolution and Culture*. Alemania: Mohr Siebeck.
- Zahavi, D. (2014). *Self and other: exploring subjectivity, empathy, and shame*. OUP Oxford.
- Zahavi, D. y Satne, G. (2015). Varieties of shared intentionality: Tomasello and classical phenomenology”, En Bell, J., Cutrofello, A. y Livingston, P. (Eds.), *Beyond the Analytic-Continental Divide: Pluralist Philosophy in the Twenty-First Century*, Routledge, London, pp. 305-325.
- Zlatev, J. (2016). Preconditions in Human Embodiment for the Evolution of Symbolic Communication. En Etzelmüller, G. y Tewes, C., (Eds.). (2016). *Embodiment in Evolution and Culture*. Alemania: Mohr Siebeck.