



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN FILOSOFÍA DE LA CIENCIA

**LA NATURALIZACIÓN NO-ELIMINATIVISTA
DEL CONOCIMIENTO:**

UN ANÁLISIS DESDE LA NORMATIVIDAD EPISTÉMICA

TESIS QUE PARA OPTAR EL GRADO DE:
DOCTOR EN FILOSOFÍA DE LA CIENCIA

PRESENTA

ALEJANDRO VÁZQUEZ DEL MERCADO HERNÁNDEZ

TUTORES PRINCIPALES

CLAUDIA LORENA GARCÍA AGUILAR (IIFs-UNAM)

EDUARDO GARCÍA-RAMÍREZ (IIFs-UNAM)

COMITÉ TUTOR: AXEL BARCELÓ ASPEITIA (IIFs-UNAM)

CIUDAD DE MÉXICO, AGOSTO, 2018



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.

A Gloria y Alejandro.

*Philosophy would render us entirely Pyrrhonian,
were not nature too strong for it.*

-DAVID HUME

CONTENIDOS

Introducción

0.1. Descripción general del proyecto.....	5
0.2. Resumen de los capítulos.....	7
0.3. Antecedentes y futuro del naturalismo epistemológico.....	16
0.4. Requisitos previos	20
0.5. Agradecimientos.....	21

1. Naturalización del conocimiento

1.1. Introducción	22
1.2. Presupuestos epistemológicos	25
1.3. El eliminativismo sobre el conocimiento	29
1.4. Naturalidad	37
1.5. Estrategias de naturalización	48
1.6. Conclusión del capítulo	52

2. Fiabilismo ecológico

2.1. Introducción	54
2.2. Conocimiento etológico	56
2.3. Argumento metodológico	61
2.4. Interpretación epistemológica del conocimiento etológico	70
2.5. Conclusión del capítulo	85

3. El conocimiento como función propia

3.1. Introducción	89
3.2. El comportamiento antipredatorio	92
3.3. Las teorías teleológicas de funciones biológicas	97
3.4. La evaluación cognitiva de funciones biológicas	104
3.5. Las funciones seleccionadas no son una condición suficiente	108
3.6. Las funciones seleccionadas no son una condición necesaria	112
3.7. Conclusión del capítulo	117

4. El conocimiento como función identificada de capacidades sistémicas

4.1. Introducción.....	119
4.2. Descripción de la tesis	121
4.3. Epistemología de la virtud	124
4.4. Dominio de evaluación	128
4.5. Unidad de evaluación y condiciones de satisfacción	132
4.6. Identificación de funciones e individuación de sistemas	137
4.7. Ejemplos de aplicación y objeciones	143
4.8. Conclusión del capítulo	152

5. Conclusión

6. Bibliografía

INTRODUCCIÓN

0.1. Descripción general del proyecto

En esta investigación llevo a cabo dos objetivos, hasta cierto punto entrelazados. Por una parte, el de realizar un trabajo de segundo orden —que podríamos llamar meta-epistemología— sobre qué elementos metodológicos, ontológicos y epistémico-normativos conlleva el proyecto de naturalizar el conocimiento proposicional¹; incluyendo por supuesto al conocimiento humano. El segundo objetivo, es llevar a cabo una evaluación crítica de los proyectos de naturalización del conocimiento —desde un considerable nivel de abstracción— así como la defensa de una alternativa naturalista que considero que no ha sido desarrollada propiamente en la literatura. El resultado no es una propuesta concreta acerca de qué es el conocimiento, dado que esto es un trabajo empírico según el naturalismo que defiendo, sino un esquema general de qué es el conocimiento.

Con ambos objetivos busco explorar el problema filosófico que acarrea la tesis de que el conocimiento es natural (usualmente expresada como que el conocimiento es una clase natural), lo que puede ser llamado un naturalismo ontológico o sustantivo, en contraste con un naturalismo metodológico (i.e. el conocimiento debe ser estudiado por medios empíricos). La intención es que este naturalismo ontológico lo sea en un sentido interesante; es decir, no sólo que los ejemplares de conocimiento que poseen los individuos descansan (e.g. "supervienen", "están fundamentados", etc....) en última instancia en propiedades naturales, sino que el conocimiento es natural de manera constitutiva. En otras palabras, el conocimiento no es una clase artificial,

¹ En adelante, sólo lo llamaré "conocimiento." Cuando hable de otros tipos de conocimiento distintos al proposicional lo aclararé específicamente.

como sería el caso si fuera una convención o construcción social, un artefacto o —como se ha defendido de manera más reciente— una institución.

Un elemento crucial de este trabajo es que se trata de una naturalización no-eliminativista del conocimiento. Es decir, al hablar de "conocimiento" no me refiero en términos vagos a lo que podríamos llamar el fenómeno epistémico: la idea de que hay maneras más adecuadas o mejores de generar representaciones sobre el medio ambiente. La intención tanto de las propuestas que analizo, como de la alternativa que propongo, es defender que el conocimiento se conserva en esta naturalización.² Se trata de que el mismo fenómeno que han estudiado los epistemólogos como un término teórico³, en su uso preteórico o folk, se conserve en esta naturalización.

Al mismo tiempo, he adoptado como parte de los desiderata que el resultado de esta naturalización no pierda el componente normativo que tiene el conocimiento. Preteóricamente, cuando realizamos evaluaciones epistémicas para atribuir conocimiento, consideramos que un agente cognitivo debe buscar obtener conocimiento respecto a meras creencias verdaderas, o que tener conocimiento es algo *prima facie* bueno (o como indica la literatura más contemporánea, valioso). Mi objetivo al mantener este desiderátum no se basa en una simple deferencia hacia nuestras nociones preteóricas, sino que considero que no podemos prescindir del trabajo evaluativo que lleva a cabo el término en nuestras prácticas epistémicas ni del trabajo teórico que realiza en nuestro intento de entenderlas. Un signo de que la normatividad epistémica en una metafísica naturalista del conocimiento no sólo es deseable sino ineludible, es que el componente normativo aparece en todos los trabajos que han intentado defender un naturalismo ontológico.

² Se puede entender como un realismo sobre el conocimiento, de manera análoga al realismo que se ha sostenido en filosofía de las ciencias cognitivas sobre la creencia (e.g. Fodor, 1975).

³ Aunque las características finales que resulte tener el conocimiento puedan diferir fuertemente de los diversos análisis que han sido propuestos a priori, lo cual no es sorprendente desde una perspectiva naturalista.

Dados estos estreñimientos, quienes defendemos un naturalismo no-eliminativista acerca del conocimiento tenemos grandes retos conceptuales por delante. Podríamos decir que, por ello, el proyecto forma parte de una etapa fundacional, donde lo que se ha buscado en la literatura es discutir cuestiones relacionadas con los fundamentos de un naturalismo sobre el conocimiento. Mi investigación forma parte de este momento teórico, y por ello me concentro de manera más abstracta en la línea general que debe seguir una naturalización del conocimiento. Sin embargo, parte de las consecuencias de un naturalismo de este tipo es que posteriormente será necesario un trabajo más granular, llenando los detalles sobre cómo se lleva a cabo el conocimiento —tanto en humanos como en otros animales— y que este trabajo se dará en colaboración con las ciencias cognitivas y otras ciencias, como la biología.

0.2 Resumen de los capítulos

El **primer capítulo** trata sobre la naturalización del conocimiento en general, es decir, sin tener en mente una propuesta de naturalización en particular. El objetivo es dar algunas nociones previas que puedan ayudar en el trabajo posterior y en establecer algunos principios metodológicos que sirvan como guía. Comienzo por caracterizar al naturalismo ontológico, distinguiéndolo del naturalismo metodológico, y sostengo que, a pesar de no tratarse de asuntos ortogonales, tienen una relativa independencia. La investigación se concentrará sobre un naturalismo ontológico, particularmente acerca del conocimiento.

El objeto de estudio se delimita al conocimiento proposicional, el que figura en enunciados del tipo "S sabe que P." También señalo algunas desiderata generales, como características que delimitan al fenómeno del conocimiento, tales como que tiene un elemento representacional (usualmente caracterizado como una creencia), un componente fáctico (que dicha creencia sea

verdadera), una relación más o menos robusta con la verdad de dicha creencia de modo que no sea gratuita, así como funciones evaluativas y normativas. También hago explícita una prioridad acerca de la creencia en el estudio del conocimiento. Es decir, el conocimiento es una propiedad que tienen ciertas creencias (lo cual no implica que pueda llevarse a cabo un análisis del conocimiento en términos de creencia verdadera más otras condiciones). Es decir, rechazo un enfoque primitivista como el de Williamson, que han adoptado algunos naturalistas como Jennifer Nagel. Asimismo, hago explícita la decisión —o quizá podría decirse, el prejuicio— de tomar al conocimiento como un fenómeno común y además que puede atribuirse a infantes y animales no-humanos.

En el siguiente apartado describo el eliminativismo y el deflacionismo sobre el conocimiento, distinguiéndolo del eliminativismo global, y doy algunas razones para rechazar esta postura e intentar conservar al conocimiento dentro de una epistemología naturalista. Retomo brevemente el naturalismo epistemológico de Quine, para caracterizarlo como un deflacionismo sobre el conocimiento y también expongo la teoría pragmática de evaluación cognitiva de Stich, como un ejemplo de eliminativismo global.

Posteriormente me concentro sobre la tesis de que el conocimiento es natural, mostrando algunas maneras como se puede caracterizar de manera más precisa. Dentro de estas caracterizaciones, descarto la noción de clase natural proveniente del esencialismo a partir de Kripke y considero que la noción de clase natural propuesta por Richard Boyd es más adecuada para el problema en cuestión. Dicha noción considera a las clases naturales como cúmulos de propiedades causalmente homeostáticas. También mantengo abierta la posibilidad de considerar la naturalidad del conocimiento como una cuestión de grado, al estilo de Lewis (1983). Si el conocimiento es una propiedad que tienen ciertas creencias, la tesis tomará la forma de que dicha propiedad tiene un grado de naturalidad cercano al de otros términos biológicos. En este mismo apartado trato el

problema sobre si la naturalización puede ser exitosa si el resultado es que el conocimiento es una clase disyuntiva (como el "jade", que es un término que se usa indistintamente para dos minerales distintos con apariencia similar).

Finalmente, trato el problema de las estrategias de naturalización. La primera estrategia que considero es la de localización, inspirada en la metafísica estilo Canberra de la Universidad Nacional de Australia. Esta estrategia consiste en darle peso a nuestra noción preteórica por naturalizar —el conocimiento, en este caso— y buscar sus realizadores (en una reconstrucción racional del proceso, esto aparecería como una oración de Ramsey). La estrategia la descarto por no considerar que el resultado garantice la naturalización del conocimiento, en el caso de que los realizadores sean múltiples. La siguiente estrategia que considero es la de fijación de referente, defendida por Kornblith. La idea es que, dado que las clases naturales tienen una semántica externista, aquello que identificamos preteóricamente como conocimiento designa algo rígidamente (en caso de una naturalización exitosa); a pesar de que posteriormente se descubra que el referente no tiene las notas características que se usaron para la designación original. En este caso, debemos abandonar la noción preteórica lo antes posible y comenzar a estudiar directamente el referente, sin tener en cuenta qué tanto cumple con las características que sirvieron para designar el término. La tercera y última estrategia que describo es la de acomodación (parcialmente inspirada en el proceso de "acomodación" de Boyd para encontrar clases naturales, pero que difiere en aspectos importantes). Esta es la estrategia que considero adecuada para llevar a cabo la naturalización de un término epistémico como el conocimiento. En este caso no nos centramos en el contenido preteórico de la noción que buscamos naturalizar, sino en el trabajo teórico que realiza (en este caso, tanto inductivo para predecir conducta, como normativo), tomando a la epistemología como si se tratara del lenguaje teórico de una ciencia. Esto conlleva el problema de que distintas teorías epistemológicas describen al conocimiento de maneras distintas.

En este caso podemos comparar distintos candidatos para ser los realizadores del conocimiento con distintas teorías epistemológicas y modificándolas de acuerdo con lo que aprendamos sobre el referente; todo esto cuidando que el producto resultante lleve a cabo el trabajo teórico de manera adecuada.

En el **segundo capítulo**, describo y reconstruyo la propuesta de naturalización que llamo fiabilismo ecológico, la cual ha sido propuesta y defendida por Hilary Kornblith en las últimas dos décadas. La intención es evaluarla críticamente y finalmente rechazarla. Lo que propone Kornblith es que la noción de conocimiento de la etología cognitiva es el referente tanto de nuestro conocimiento preteórica como aquello sobre lo que han teorizado los epistemólogos. Dado que Kornblith utiliza como uno de sus argumentos la indispensabilidad, que tiene que ver con el uso necesario de determinados términos por una ciencia y debido a que existen críticas al carácter científico del conocimiento etológico, realizo un breve recuento de los orígenes de dicha disciplina en su contexto intelectual como un modo de dar a conocer la raíz de sus compromisos metodológicos. Posteriormente, expongo la noción de conocimiento utilizada en la etología, centrándome en los ejemplos citados por el propio Kornblith.

Después, reconstruyo el apoyo que da Kornblith para tomar al conocimiento etológico como referente del conocimiento —incluyendo el caso particular del conocimiento humano— bajo la forma de lo que denomino su argumento metodológico. En esta reconstrucción expongo tanto lo que considero que son los presupuestos del argumento, como las premisas, examinando la plausibilidad de ambos. Los presupuestos tienen que ver con el externismo semántico de las clases naturales, con los compromisos ontológicos con los términos indispensables, el estatus de la etología cognitiva como ciencia madura y el hecho de que el conocimiento realmente sea un término de la etología cognitiva de una manera no-trivial. En este análisis de los presupuestos y de las premisas del argumento metodológico retomo algunos de los elementos del primer capítulo

acerca de las estrategias de naturalización, para mostrar cómo afectan la plausibilidad de las premisas.

Lo siguiente es una explicitación de la manera como Kornblith integra el conocimiento etológico a la epistemología reconstruyéndolo en términos filosóficos más familiares. A esta reconstrucción la llamo fiabilismo ecológico. La parte fiabilista tiene que ver con que los mecanismos generadores de conocimiento producen creencias verdaderas de manera fiable. Aquí interpreto la fiabilidad en un sentido estricto, como una tasa de creencias verdaderas sobre falsas. El aspecto fiabilista de la postura se ve claramente en la medida en que el fiabilismo ecológico es una respuesta a problemas del fiabilismo de proceso (o método) tradicional, particularmente el problema de la generalidad. La parte ecológica tiene que ver con que el ambiente cognitivo *para* el cual dicho mecanismo fue seleccionado es el que determina la manera de individuar los procesos. Abordo posteriormente la tesis de que el conocimiento, entendido como fiabilismo ecológico, puede explicar la totalidad del fenómeno del conocimiento humano; lo cual implica rechazar una tesis como la de Sosa de que además del conocimiento animal hay un conocimiento de un segundo nivel (e.g. un conocimiento reflexivo).

Para finalizar el capítulo, llevo a cabo una crítica al fiabilismo ecológico. Por una parte, el requisito de fiabilidad parece incompatible con los mecanismos propuestos dentro de las ciencias cognitivas que son relevantes para el estudio del conocimiento, incluyendo a la etología cognitiva. Esto es debido a que no son raros los mecanismos que generan una gran cantidad de creencias falsas, como sucede con los falsos positivos de los mecanismos de detección de depredadores. Ante esto, el fiabilista ecológico tiene un dilema. Por una parte, puede insistir en que sólo los mecanismos que fueron seleccionados por ser fiables producen conocimiento, una distinción que no se encuentra en la literatura de modo que la naturalidad del criterio tendría que justificarse. La otra alternativa es retractar el requisito de fiabilidad, o sustituirla por una que esté definida en términos

evolutivos para hacer compatibles los elementos fiabilistas de la teoría con la parte ecológica. Pero esto hace colapsar al fiabilismo ecológico en el conocimiento como función propia (o al menos no queda claro en qué se distinguirían ambas posturas). Sostengo que este dilema, y algunos otros problemas, deben llevar al abandono del fiabilismo ecológico.

En el **tercer capítulo** expongo la teoría del conocimiento como función propia, también llamada teleológica, para señalar sus problemas y argumentar que tampoco es una manera adecuada de naturalizar el conocimiento. La tesis indica que lo que distingue a las creencias verdaderas que son conocimiento de las que no lo son, es que fueron producidas por mecanismos cuya función biológica es producir creencias del tipo relevante al que pertenece dicha creencia; y que lo que determina las funciones biológicas son propiedades evolutivas relacionadas con la selección natural. A diferencia de lo que hago en el segundo capítulo con Kornblith, no me centro en una versión particular de la tesis. La tesis que considero, el conocimiento como función propia, dice que una creencia verdadera de un cierto tipo es conocimiento si y sólo si fue producida por un conjunto de mecanismos generadores cuya función es producir creencias de dicho tipo. El aspecto central tiene que ver con la noción de función utilizada, que se trata más adelante en este mismo capítulo.

Antes de adentrarme en ello, doy una descripción breve de la etapa de representación y clasificación en el comportamiento antipredatorio para mostrar un ejemplo de algo que podría ser conocimiento bajo una teoría teleológica. Algo notorio es que no hay una identificación con un término teórico "conocimiento", como propone Kornblith, sino que se trata de una naturalización que permite un pluralismo teórico (distintas disciplinas, incluyendo a la etología cognitiva, pero también a la psicología comparada, etc.). Trato también el problema de cómo las representaciones de animales no-humanos con altas capacidades cognitivas pueden ser descritas como creencias, a pesar de no estar constituidas por conceptos con un nivel de generalidad tan grande como el de

los conceptos humanos (no a partir de nociones empíricas de conceptos psicológicos en humanos, sino desde el punto de vista de las teorías filosóficas tradicionales, como la semántica de mundos posibles).

Posteriormente, hago un breve recuento histórico de las teorías teleológicas de función biológica, mismas que atribuyen funciones a los rasgos a partir de nociones selectivas — desde sus orígenes en Wright— para exponer las distintas variantes de la tesis. En particular me centro en la versión etiológica de la tesis, que atribuye funciones a partir del origen histórico del rasgo, como ha sido defendida por exponentes como Millikan y Neander. Muestro algunos problemas relacionados con la asignación de funciones a partir de distintas maneras como se pueden interpretar los requisitos de haber sido seleccionado, dividiendo retomando una división de las teorías etiológicas en fuertes (seleccionado *para*) y débiles (seleccionado *por*), mostrando los problemas que presentan ambas interpretaciones para asignar funciones y para explicar los casos de disfunciones.

Una vez hecho esta recapitulación sobre las teorías teleológicas de funciones biológicas, dedico un apartado a los retos que conlleva identificar a las evaluaciones cognitivas como funciones biológicas (principalmente del aparato cognitivo⁴). Posteriormente dedico dos apartados a criticar la tesis del conocimiento como función propia, uno para mostrar que no es una condición suficiente para el conocimiento y otro que no es una condición necesaria.⁵

Las funciones propias no son condiciones necesarias para el conocimiento debido a que existen mecanismos cognitivos que no son producto de la selección natural (como en el caso de algunos

⁴ Entendido en sentido amplio, por ejemplo las glándulas suprarrenales —que producen adrenalina— pueden formar parte de ciertos procesos cognitivos.

⁵ El proyecto naturalista no está comprometido con que el estudio del conocimiento sea un análisis por condiciones necesarias y suficientes (Cfr. e.g. Kornblith 2002, cap. 1 y Millikan, 1984). Uso este lenguaje de manera un tanto laxa, refiriéndome a que hay una plétora de casos que muestran que las teorías teleológicas no capturan de manera adecuada el fenómeno. Agradezco a Dubian Cañas por hacerme notar esto.

individuos neurodivergentes) que cumplen con los desiderata del conocimiento sin haber sido seleccionados y en ocasiones sin tener un valor adaptativo. También menciono brevemente cómo el neuroconstructivismo y la ontogénesis pueden ser razones para no considerar a las funciones propias como una condición necesaria para el conocimiento. Por otra parte, considero que tampoco son una condición suficiente, dado que la selección natural opera sobre la aptitud y que —como se ha sostenido ampliamente en la literatura— la aptitud del aparato cognitivo no necesariamente coincide con valores estrictamente epistémicos (a manera de ejemplo, para la aptitud también entran en consideración requerimientos energéticos. De manera general, señalo los problemas de considerar que la aptitud necesariamente coincide con rasgos adaptativos. Este capítulo, por tanto, muestra las ventajas de considerar al conocimiento como función, pero a la vez pone en evidencia los problemas de asignar las funciones biológicas en general —y concretamente las funciones cognitivas— a partir de las teorías teleológicas.

Finalmente, en el **cuarto capítulo** expongo y defiendo la tesis del conocimiento como función, entendiendo "función" a partir de una teoría de capacidades sistémicas derivada del trabajo seminal de Cummins. En dichas teorías, la función se entiende como una contribución causal que realiza un subcomponente de un sistema del que forma parte para llevar a cabo una determinada acción, con el requisito de que dicho sistema debe estar organizado jerárquicamente para que sea teóricamente fructífero realizar esta conceptualización en sistemas y subsistemas. A diferencia de la interpretación recibida del trabajo de Cummins como atribucionista, donde las funciones se asignan libremente, introduzco el término de funciones identificadas para resaltar el hecho de que las funciones se descubren y no se asignan. Las funciones son objetivas, debido a que una función implica que existe una organización jerárquica como la anteriormente mencionada y las descubrimos gracias a lo que ha sido denominado en la literatura como destacamiento funcional [functional salience].

Como antecedente, hago una recapitulación de las teorías de la virtud fiabilistas, dado que el resultado pretendido es algo muy parecido a una naturalización de dicha noción (entendiendo a la fiabilidad de un modo amplio). Una vez hecho este recuento, comienzo a elaborar sobre los elementos que considero que deben tener las evaluaciones cognitivas: un dominio de evaluación, unidad de evaluación y condiciones de satisfacción, incluyendo una noción muy austera de agencia epistémica que denomino cuasi-agencia.

Una vez hecho esto, muestro cómo los mismos elementos de las evaluaciones cognitivas pueden servir para identificar funciones cognitivas. Para esta parte me apoyo en el trabajo que se ha llevado a cabo en la filosofía de la ciencia mecanicista, especialmente mostrando cómo la continuidad productiva es un elemento de individuación de los sistemas cognitivos. En el caso que me interesa, la capacidad de producir creencias verdaderas y adecuadas en determinados contextos de manera regular, es lo que permite identificar a un sistema como generador de conocimiento (y por tanto a las creencias resultantes como ejemplares de conocimiento). El resultado es que guiados por intereses epistemológicos (i.e. sobre cómo los organismos se representan adecuadamente el ambiente en el que viven) podemos *descubrir* sistemas que en parte están constituidos como tales por realizar dicha función (posiblemente entre otras). Este paralelismo entre las evaluaciones cognitivas y las evaluaciones de funciones cognitivas, es lo que permite tener una teoría del conocimiento a la vez naturalista y que cumpla con las características normativas descritas.

En lo que resta del capítulo doy un breve ejemplo de cómo se puede aplicar la teoría del conocimiento como función de capacidades sistémicas, tomando el caso del conocimiento nuclear [core knowledge] descrito por Carey y Spelke. Se trata de creencias generadas por sistemas innatos de dominio específico, que se encuentran en infantes prelingüísticos, grandes simios y universales desde el punto de vista antropológico. Posteriormente pongo sobre la mesa las objeciones que considero más fuertes en contra de la teoría, mostrando algunas vías posibles para resolver cada

una de ellas. Haciendo un balance general, concluyo que el conocimiento como función de capacidades sistémicas es la manera más plausible de naturalizar al conocimiento, manteniendo sus características más importantes.

0.3. Antecedentes y futuro del naturalismo epistemológico

El naturalismo epistemológico contemporáneo se remonta a Quine, quien en "Epistemology Naturalized" (1969a) sugiere reformular la epistemología como un uso particular de la psicología y otras ciencias naturales. De acuerdo con Quine, el papel de la epistemología es —por una parte— explicar cómo de austeros insumos podemos llegar a grandes productos como teorías científicas, y por otra, realizar una evaluación crítica de nuestras maneras de conocer. Por razones ligadas con trabajos anteriores (1951), concluye que es imposible llevar a cabo estos ideales tal como se intentó por buena parte de la filosofía de la primera mitad del siglo 20 por autores como Carnap y Hempel que consistía en fundamentar a las teorías científicas —desde un punto de vista semántico⁶ y epistémico— a partir de datos sensoriales y principios lógicos. Esto era una fundamentación independiente, que no se apoya en el cuerpo de conocimientos científicos existentes.

Quine sostiene que una fundamentación de carácter tan ambicioso es imposible, pero la psicología es capaz de iluminar tanto la relación entre los insumos perceptuales y productos teóricos, como de criticar y evaluar nuestros procesos epistémicos. Esto conlleva un problema de aparente circularidad, pues a su vez la psicología —y otras ciencias auxiliares⁷— a su vez forman parte del cuerpo de conocimientos de la ciencia. La propuesta de Quine es aprender a vivir con la circularidad, mostrando que por una parte no es viciosa y por otra, es inescapable.

⁶ Particularmente, garantizar que los términos teóricos tengan significatividad empírica.

⁷ Por ejemplo, la óptica puede explicar cómo funcionan algunos instrumentos observacionales.

El trabajo de Quine suscitó respuestas que constituyeron un primer momento del naturalismo epistemológico contemporáneo. Una de las contribuciones de la literatura posterior consiste en refinar las distintas tesis del naturalismo epistemológico, como separando la tesis sustantiva (lo epistémico descansa en última instancia en propiedades naturales) de la metodológica (la epistemología debe llevarse a cabo con ciencias empíricas). Por otra parte, surgen textos críticos que cuestionan que la posibilidad de una epistemología naturalizada (Putnam, 1982), o que ésta pueda ser capaz de llevar a cabo la dimensión normativa de la epistemología tradicional (Kim, 1988); así como reacciones de apoyo moderado (Goldman 1994) e intentos de defender a Quine de la acusación de circularidad viciosa (Kitcher, 1992). Asimismo, la discusión se generaliza. Mientras que el texto seminal de Quine fue escrito en el contexto de realizar una epistemología fundacional para la ciencia, las discusiones posteriores acerca del naturalismo epistémico incluyen también una epistemología acerca de los estados epistémicos de los sujetos en situaciones ordinarias. En palabras de Kitcher (1992, p. 56):

Las epistemologías naturalistas confrontan una gran cantidad de cuestiones tradicionales: ¿Qué es el conocimiento? ¿Qué tipos de conocimiento son posibles [...]? ¿Qué métodos debemos usar para alcanzar el conocimiento o por lo menos para mejorar la calidad epistémica de nuestras creencias?

De este modo, algunos epistemólogos tradicionales comienzan a incluir elementos metodológicos en su trabajo, como Goldman (1986) en sus trabajos fiabilistas; o quienes buscan naturalizar la idea de virtud epistémica⁸ (ver la compilación *Naturalizing Epistemic Virtue*, Fairweather, 2014). Más cerca de nuestro tiempo, hay filósofos que han adoptado una postura más fuerte. Jennifer Nagel (2008) es un buen ejemplo de esto, pues no parte de una discusión acerca del naturalismo epistemológico, sino que analiza directamente los trabajos de psicología que tratan a las atribuciones de conocimiento y otros estados epistémicos, buscando mostrar cómo dichas

⁸ Como se podrá ver en el mismo capítulo, suele haber elementos fiabilistas en las teorías de la virtud epistémica.

investigaciones resuelven algunas de las preguntas filosóficas acerca del fenómeno epistémico. Cabe mencionar que también existe el caso de Millikan en "Naturalist reflections on knowledge" (1984), quien a pesar de no dedicarse a la epistemología busca incluir al conocimiento dentro de una naturalización general acerca de los estados mentales, las funciones y el lenguaje.

De manera paralela y sin relación con esta literatura, algunos psicólogos cognitivos que estudian el razonamiento humano comienzan a interesarse por el aspecto normativo de la racionalidad. Los trabajos de Kahneman y Tversky (e.g. 1982) muestran los sesgos cognitivos de los razonadores humanos y los consideran fallos de racionalidad; mientras que otros investigadores como Gigerenzer (2010) y Todd realizan nuevos experimentos que ponen en contexto dichos resultados y toman la moraleja distinta de que la racionalidad es ecológica; es decir, que no es independiente del contexto y que para evaluarla hay que considerar los ambientes cognitivos adecuados. Esto incluye la manera en que se presenta la información; pues, por ejemplo, el éxito del razonamiento probabilístico depende de si a los sujetos se les presentan porcentajes o tasas (Cfr. Stanovich y West, 1998; Kahneman y Tversky, 1986). Otro ejemplo se puede encontrar en el trabajo de Cosmides y Tooby (1992), donde se muestra cómo el razonamiento con implicaciones materiales es exitoso cuando el contenido de las implicaciones tiene que ver con reglas sociales (Cosmides y Tooby, 1992) pero no en abstracto ni en otros contextos, lo que lleva a pensar que las competencias de razonamiento no son independientes del dominio.

La literatura anterior proveniente de la psicología del razonamiento constituye los inicios de lo que posteriormente será llamado el debate de racionalidad (Stein, 1996) —o de manera más dramática, la guerra de racionalidad. Este debate, al cual se suman los filósofos, incluye diversas preguntas: ¿qué se puede extraer de la evidencia experimental sobre sesgos cognitivos?, ¿es coherente defender la tesis de que los seres humanos somos irracionales?, ¿la evidencia empírica es compatible con el pluralismo epistémico (Eraña, 2003)? La discusión puramente filosófica acerca

de la naturalización de la epistemología se encontró con las preocupaciones normativas surgidas de la investigación empírica.

Los filósofos que iniciaron con el naturalismo contemporáneo tenían un interés de salir de las ideas intelectualistas dominantes en la epistemología tradicional, mientras que los psicólogos participantes en el debate de racionalidad deseaban ir más allá de lo puramente descriptivo. El resultado ha sido un intercambio fructífero, que ha mostrado que las preguntas sobre el naturalismo epistemológico tienen más sentido cuando se piensan a la luz de las investigaciones empíricas existentes, y que a su vez las implicaciones de éstas resultan más claras dentro de un marco de discusión con conceptos bien delimitados con la claridad que la filosofía puede proveer. La manera en que se han abordado estas cuestiones, muestran que ya existe una tradición de naturalismo epistemológico —aunque claramente no es monolítica, sino incluso algo caótica— que continuará madurando en el futuro previsible.

No obstante, estos son avances en la cuestión metodológica, mientras que la cuestión ontológica del naturalismo ha sido mucho menos atendida. El trabajo de Kornblith (2002) ha sido crucial para poner en primer plano la cuestión sobre el estatuto ontológico de las nociones epistémicas, y también —como se puede ver claramente— la inspiración más directa para este trabajo. ¿Son naturales el conocimiento, la justificación, la racionalidad, de la acreditación? ¿El naturalismo metodológico conlleva eliminar estos términos epistémicos para sustituirlos con otros con mayor significatividad empírica y que correspondan de manera más adecuada a la realidad de los agentes cognitivos? Si el conocimiento puede sobrevivir dentro de una visión naturalista de la epistemología, ¿qué es y dónde podemos esperar encontrarlo?

A mi modo de ver, estas preguntas son aún más difíciles y más urgentes que las cuestiones metodológicas. Para que el naturalismo epistemológico contemporáneo pase de este primer

encuentro fructífero que describí anteriormente, a un segundo momento de claridad —aunque no necesariamente unidad— teórica. Hace falta una mayor delimitación acerca de los objetos de estudio, así como una claridad mínima acerca de lo que son, y cómo pueden ser abordados en sus distintos aspectos tanto por disciplinas filosóficas, como por las ciencias naturales y sociales.

El conocimiento es un elemento paradigmático de la epistemología. No sólo ha sido objeto de reflexión durante siglos, sino que incluso que es la noción central de la epistemología contemporánea (junto con la justificación). La pregunta acerca del estatus ontológico del conocimiento debe ser un primer paso para comenzar a aclarar estas cuestiones, cuya solución permitirá llevar el desarrollo del proyecto epistemológico naturalista hacia nuevas direcciones.

0.4. Requisitos previos

Pensando en el improbable caso de que algún lector cuya formación no sea filosófica se encuentre con esta tesis, ofrezco a continuación una guía breve del trasfondo teórico. A pesar de que este trabajo no profundiza en las discusiones de epistemología tradicional, presupone algunas nociones que pueden suplementarse en un trabajo introductorio como el de Pritchard (2014); y aunque dedico un espacio considerable a presentar y reconstruir la postura Kornblith (2002), el capítulo 1 del trabajo citado es una excelente introducción para las motivaciones del naturalismo ontológico desde una perspectiva metodológica. También pueden ser de utilidad textos panorámicos sobre representación mental (Cummins, 1991 y Fodor, 1985) y reduccionismo (Esfeld y Sachse, 2004). Asimismo, este trabajo utiliza algunas nociones metafísicas —como *necesidad* o *esencia*— que pueden consultarse en los capítulos relevantes del libro de texto de Lowe (2002) (aunque en la mayoría de los casos una comprensión intuitiva de los términos será suficiente).

En cuanto a la parte empírica, el texto presupone una familiaridad mínima con la teoría evolutiva y con el estudio de la cognición animal. Dado que los objetivos de esta investigación son fundamentalmente filosóficos, la discusión no se suele adentrar en cuestiones técnicas, pero los lectores que lo crean necesario pueden recurrir al trabajo de divulgación de Mayr (2002) y a los libros de texto respectivos de Bermudez (2014) y Andrews (2015).

0.5 Agradecimientos

Este trabajo fue realizado gracias al apoyo de una beca CONACyT del Programa Nacional de Posgrados de Calidad (2013-2017) y del proyecto PAPIIT IA401717 "Pluralismo y normatividad en lógica y matemáticas" de la Universidad Nacional Autónoma de México, a cargo de Cristian Gutiérrez. Le agradezco a los investigadores miembros de mi comité tutorial, Claudia Lorena García, Eduardo García-Ramírez, y Axel Barceló, por su guía constante y su dedicación a lo largo de esta investigación. También agradezco por sus sugerencias y correcciones a Jonatan García y Ángeles Eraña, quien también me apoyó durante la elaboración de la propuesta original. Quisiera dar las gracias también a la coordinación del Posgrado en Filosofía de la Ciencia y de manera muy especial a Marisela López a cargo de la Secretaría Académica. Este espacio es muy breve para enlistar a todos los colegas cuyas observaciones y retroalimentación me ayudaron para entender mejor los problemas sobre los que trata esta tesis, pero quisiera mencionar en particular a Moisés Macías Bustos, Dubian Cañas, María Martínez Ordaz, David Fajardo, Carlos Romero, Paloma Hernández, Max Martínez y Hilary Kornblith. Finalmente agradezco a mis padres, a mi hermana Montserrat, a todos mis amigos y a Ana por su apoyo incondicional durante la elaboración de este trabajo.

CAPÍTULO 1: NATURALIZACIÓN DEL CONOCIMIENTO

1.1. Introducción

Al hablar de naturalismo epistemológico, lo primero que viene a la mente es la amplia literatura sobre una discusión metodológica: la pregunta de cuál debe ser el papel de las ciencias empíricas en la epistemología. Las respuestas oscilan desde una sustitución radical hasta la afirmación de que éstas no son filosóficamente relevantes, con varias tesis intermedias sobre cómo debería ser una colaboración entre ambas. Un asunto que va de la mano es el vocabulario teórico de la epistemología. En la medida en que una postura sea naturalista tendrá un compromiso menor por conservarlo y estará dispuesta en mayor medida a sustituirlo por los términos de las ciencias que realicen la mayoría del trabajo teórico o incluso que sean candidatos para reemplazarlo. Algo similar sucede con el grado en que se habrán de conservar los problemas epistemológicos tradicionales, al menos en la manera como han sido formulados en la literatura filosófica. Se trata de un conjunto de debates sumamente interesantes que están lejos de ser resueltos.

No obstante, en esta investigación trataré un problema distinto que también versa sobre la naturalización, pero a partir de la metafísica del conocimiento. La pregunta, a grandes rasgos, es si el conocimiento es natural. Se trata de una tesis que, si no se hacen mayores precisiones, resulta difícil de examinar por su vaguedad. El objetivo primordial de este capítulo es formularla de manera más precisa, explicitando lo que conlleva.

Desafortunadamente, los problemas metafísicos acerca del conocimiento (y otras nociones epistémicas) han sido enormemente descuidados: en términos generales ha comenzado a profundizarse sobre ello apenas en las últimas dos décadas. A pesar de que en la epistemología — como en cualquier disciplina— las teorías tienen distintos compromisos ontológicos, éstos apenas

han comenzado a examinarse críticamente en la meta-epistemología, un área de estudio que aún está en proceso de definirse.⁹

El propósito de lo anterior ha sido distinguir los problemas metodológicos de los ontológicos. Por supuesto, ambas familias de problemas no son completamente independientes entre sí, ni tampoco son independientes de los problemas centrales de la epistemología (lo cual es esperable, dado que una tesis metafísica puede tener consecuencias metodológicas, epistemológicas, etc.). De hecho, es común encontrar ambos aspectos entremezclados en la literatura. Sin embargo, una señal de cierto grado de independencia de ambas cuestiones es que se ha defendido un espectro de posturas. Con el objetivo de dar algunos ejemplos, tenemos esta matriz:

	Clase artificial	Clase natural
Metodología tradicional ¹⁰	Craig (1999)	Ahlstrom (2008) ¹¹
Metodología empírica	May et al. (2010) ¹²	Kornblith (2002, 2007) ¹³
Metodología mixta	Goldman (2007)	Alston ¹⁴ (1995)

⁹ Un ejemplo de este cambio puede verse en Goldman (2007), quien ahora defiende explícitamente la idea de que el conocimiento es una clase artificial, lo cual era solamente un presupuesto no declarado en sus trabajos más tempranos.

¹⁰ La metodología de Craig no consiste en el análisis conceptual típico, sino una reconstrucción ideal a partir de pensar en las necesidades de una comunidad epistémica en estado natural (i.e. sin nociones epistémico normativas) y en las prácticas que iría creando por necesidad. No obstante, a pesar de ser una metodología atípica la considero “tradicional” en tanto que lo empírico no juega un papel en dicha reconstrucción.

¹¹ La metodología tradicional es factible en la medida en que nuestro concepto de conocimiento coincide con las características que realmente tiene el conocimiento, lo cual no es necesario, pero de acuerdo con Ahlstrom es mínimamente plausible.

¹² La metodología propuesta es la filosofía experimental, que a pesar de ser empírica no sería considerado por todos como un naturalismo metodológico; dado que se trata de una agregación estadística de las intuiciones epistémicas de distintas poblaciones obtenidas por encuestas. Algunos experimentalistas usan la evidencia para rechazar la utilidad de las intuiciones, pero otros (como May), utilizan la metodología para conocer las intuiciones (programa positivo). Para un recuento general de la metodología experimental aplicada a epistemología, ver Weinberg (2010).

¹³ Kornblith (1980) defiende que la semántica de las clases artificiales es también externista, de lo cual se desprende la consecuencia de que la metodología tradicional sería impropcedente aún en el caso de que el conocimiento no fuera una clase natural

¹⁴ Alston ensaya la postura de un fiabilista de método que considera que los métodos se individualizan por procesos que constituyen clases naturales, para posteriormente criticarla y rechazarla.

Un ejemplo donde se puede ver la independencia parcial del problema ontológico y del problema metodológico es el trabajo citado de Ahlstrom (2008). Sostener un naturalismo ontológico no lo obliga a aceptar una metodología empírica, pero necesita postular una adecuación de los conceptos epistemológicos preteóricos con sus referentes, lo cual permite llevar a cabo una metodología tradicional de análisis a partir de intuiciones semánticas. También puede ser útil un ejemplo de la independencia de los problemas ontológico y metodológico con respecto a las cuestiones epistemológicas. Si bien el naturalismo ontológico y metodológico suelen estar asociados a una epistemología de corte externista, Lepock (2014) sugiere incluir elementos de creencias de segundo orden (como la reflexión) analizados a partir de la literatura empírica sobre metacognición. Estos dos ejemplos ilustran cómo a pesar de que el naturalismo ontológico se lleva mejor con un naturalismo metodológico, la adopción de una postura no obliga a sostener la otra.

Por lo anterior, considero que resultará más fructífera la estrategia de concentrarse en la pregunta sobre la naturalidad del conocimiento y ver sobre la marcha las relaciones que puedan existir con otros problemas. El objetivo no es presentar una meta-epistemología completa, ni siquiera una teoría global acerca del estatus metafísico del conocimiento: la cuestión de la naturalidad o artificialidad del mismo es sólo una pregunta metafísica entre muchas otras que se pueden hacer. Pero dado que la pregunta sobre si éste es natural se trata de una cuestión central, al responderla será posible ofrecer algunos avances en diversos puntos del problema metodológico. A partir de este punto, a lo largo de este trabajo cuando hable de "naturalismo" sin ninguna calificación, estará implícito que me refiero al naturalismo con respecto al problema ontológico.

Además de la distinción entre la discusión metafísica y la metodológica, vale la pena hacer una aclaración acerca de dos niveles de preguntas acerca del conocimiento, el epistemológico y el meta-epistemológico. Las discusiones metodológicas se suelen clasificar dentro de la meta-epistemología, pero ésta también incluye aspectos relacionados con la constitución de los objetos

epistemológicos. De este modo, entiendo a la epistemología sobre el conocimiento como la investigación acerca de las condiciones para que algo sea conocimiento,¹⁵ mientras que los problemas meta-epistemológicos acerca de éste tienen que ver con ciertos aspectos —metafísicos, lingüísticos, mentales— que no entran en una teoría en el nivel epistemológico. Esto puede verse de manera análoga en la distinción entre ética (concretamente, teoría moral) y meta-ética, donde la primera se centra en las condiciones de adecuación para que un acto o un agente sea considerado bueno, mientras que la segunda se concentra en debates metafísicos, lingüísticos y psicológicos relacionados con el estatuto de aquello que se califica como moralmente bueno.¹⁶ Del mismo modo como sucede en la ética, las cuestiones meta-epistemológicas no son completamente independientes de las posturas epistemológicas, pero vale la pena mantener la distinción en la medida de lo posible para facilitar el trabajo.

1.2. Presupuestos epistemológicos

En esta sección daré algunos presupuestos generales acerca del conocimiento con los cuales será posible llevar a cabo una investigación acerca de su estatus metafísico como se suele realizar en otros trabajos (e.g. Williamson, 2000). No sería deseable limitar este trabajo a una sola teoría o familia de teorías —aunque en ocasiones será necesario examinar alguna tesis metafísica desde distintas caracterizaciones del conocimiento— pero resultaría impráctico no presuponer nada. Por ello considero que lo mejor es hacer explícitos algunas acotaciones que son ampliamente compartidas en la literatura, aunque sea de manera inicial.

Adoptaré los siguientes presupuestos epistemológicos:

¹⁵ Estas condiciones no tienen que tener que ser un análisis por condiciones necesarias y suficientes, sino que puede ser otro tipo de elucidación.

¹⁶ Otro ejemplo puede verse en el trabajo de Tarski (1956) sobre la verdad. Éste propone condiciones de adecuación para que un enunciado sea verdadero, pero éstas son compatibles con diversas teorías sobre qué es la verdad; como por ejemplo, el deflacionismo y el correspondentismo.

a) El conocimiento proposicional es el objeto a estudiar. Con esto únicamente quiero indicar el tipo de conocimiento que responde a adscripciones como "S sabe que P", sin tomar una postura con respecto a si una semántica detallada acerca del contenido de los estados epistémicos utilizará proposiciones para dicho efecto.¹⁷

b) El conocimiento tiene un carácter que podemos llamar representacional (Williamson, 2000). Se trata de una formulación más general del requisito típico de que "S sabe que P" implica "S cree que P"¹⁸. Con esto no presupongo que en el conocimiento se dé algún tipo de representación específica (por ejemplo, funcional, teleosemántica), ni tampoco quiero descartar que ésta se deba caracterizar de forma disposicional. Es decir, afirmar que el conocimiento es representacional no implica aquí un representacionismo en el sentido de cuando hablamos de teorías representacionales sobre la creencia, donde hay un compromiso ontológico con la existencia de dichas representaciones como ejemplares en el aparato cognitivo de un individuo. Un ejemplo de trabajo que considera el aspecto representacional del conocimiento de forma disposicional puede ser encontrada en Armstrong (1973).

c) El conocimiento tiene un carácter "fáctico" (Williamson, 2000). Usualmente esto se expresa diciendo que "P es verdadero" es una condición necesaria de "S sabe que P", es decir, que P es el caso. Este supuesto ha sido escasamente cuestionado (Cfr. Hazlett, 2010) en la epistemología, e incluso en inteligencia artificial, si bien en algunas disciplinas de las ciencias sociales se suele utilizar el término "saber" en un sentido técnico más cercano a la creencia (e.g. el saber

¹⁷ Es decir, el conocimiento es proposicional en tanto que las atribuciones de conocimiento a nivel del lenguaje natural son sobre proposiciones, pero al aumentar el nivel de granularidad quizá no aparecerían proposiciones en la semántica.

¹⁸ A su vez, "S cree que P" junto con otras condiciones, implica conjuntamente "S sabe que P." El análisis clásico tiene otras dos condiciones: P es verdadero, la creencia de S de que P está justificada. El proyecto tradicional suele o bien proponer condiciones candidatas para agregar, o hacer una teoría de la justificación que evite los contraejemplos clásicos propuestos por Gettier (1963).

comunitario), pero este término no rescata la noción que se estudia en epistemología ni tampoco en las atribuciones de conocimiento ordinarias.

d) El conocimiento tiene un papel normativo dentro de las prácticas epistémicas, lo cual puede verse en el uso evaluativo que se hace de él.¹⁹ Decir que "S sabe que P" es una evaluación epistémica que sanciona a S, con respecto a otro agente que no sabe que P (en especial con respecto a uno que meramente cree verdaderamente que P). De aquí se deriva la necesidad de que haya en una teoría acerca del conocimiento los elementos suficientes para dar cuenta de este papel normativo.²⁰ Distintas teorías difieren acerca de qué es este elemento y en la localización del mismo, con lo cual se suelen calificar en internistas o externistas. Adoptando momentáneamente la tesis de que la justificación es parte del conocimiento, reseñaré ambas posturas de forma breve. Un modo de caracterizar a las teorías internistas es por medio de una condición acerca de que el agente tenga acceso cognitivo a aquello que justifica su creencia. La gran mayoría de las teorías internistas son doxásticas²¹: sostienen que una creencia está justificada únicamente en función del conjunto total de lo que un agente cree (Pollock y Cruz, 1999). La conjunción de ambas tesis hace que las teorías internistas (definidas a partir de la accesibilidad) identifiquen la justificación con un conjunto de razones disponibles a dicho agente. El externismo puede definirse como la negación de esta tesis, al menos hay un elemento que provee justificación (o algo análogo) que se encuentra fuera del conjunto doxástico de dicho agente. Dicho elemento usualmente hace referencia a la relación que tiene el agente epistémico con el mundo o a aspectos histórico causales sobre cómo una creencia fue adquirida o bien cómo se formó el sistema cognitivo mediante el cual fue adquirida.

¹⁹ Esta característica frecuentemente se conecta con la afirmación de que el conocimiento no puede ser "fortuito" (Pritchard, 2005) o "accidental" (Unger, 1968).

²⁰ O si se defiende que no hay dicho elemento normativo, dar cuenta de la evidencia que muestra que esto es el caso.

²¹ Pollock y Cruz (1999) defienden un internismo no-doxástico que llaman realismo directo. Es una postura atípica.

Resulta poco sorprendente que el naturalismo epistémico, tanto ontológico como metodológico, se suele formular partiendo de presupuestos externistas. Para ilustrar, supongamos que S tiene la creencia justificada C de que P, y que la justificación consiste en tener los componentes naturales $\varphi_1 \dots \varphi_n$. Siempre podemos plantear la posibilidad de un duplicado funcional S' de S, que carece de alguno de los componentes $\varphi_1 \dots \varphi_n$, pero que tiene acceso a las mismas creencias (en un sentido estrecho). La más que podría sostener un internista naturalista, es que el internismo es el caso en este mundo.

e) Un aspecto relacionado con lo anterior es el asunto del valor del conocimiento respecto de la creencia verdadera (o incluso de la creencia verdadera justificada). No supondré a lo largo de este trabajo que éste sea un valor propiamente epistémico (en contraste con meramente práctico²²), ni que este superávit de valor se cumpla en todas las instancias de conocimiento; pues existe cierta controversia con respecto a estos puntos. Sin embargo, parece plausible *prima facie* que en muchas instancias existe este valor adicional. Cabe notar que es un punto que comparten tanto naturalistas como no-naturalistas con respecto al conocimiento.

f) Un último presupuesto es que el conocimiento es un fenómeno común. Conocemos incontables proposiciones diversas tales como "mi casa está en un cuarto piso", "Beijing es la capital de China" o "2+2=4." Al presuponer la ubicuidad del conocimiento se descarta tanto el escepticismo, como la idea de que el conocimiento es un logro espectacular, como sostuvieron autores históricos como Descartes en sus *Meditaciones Metafísicas* y Hegel en la *Ciencia de la Lógica*.²³ La epistemología contemporánea suele dividir el problema de elucidar qué es el conocimiento del problema de

²² El ejemplo clásico procede de Platón (*Menón*, 70a). Un viajero rumbo la ciudad de Larisa tiene menos oportunidades de desviarse si *sabe* cuál es el camino que si meramente tiene una creencia verdadera al respecto (Cfr. Pritchard, 2010).

²³ También podría simplemente decirse que dichas nociones hiper-exigentes de conocimiento son un fenómeno distinto al que se quiere estudiar.

vindicarlo refutando al escepticismo. Esta división del trabajo permite avanzar en la investigación sobre qué es el conocimiento sin tener que entrar en otro tipo de problemas. Un punto de contención es si es correcto hablar de conocimiento y otros estados epistémicos en animales no-humanos. Esto es una tesis que casi invariablemente acepta el naturalista sobre el conocimiento, pero también una gran cantidad de externistas no-naturalistas como Goldman (2005).²⁴

Los presupuestos anteriores delimitan un marco muy general dentro del cual es posible plantear una discusión sobre la naturalidad del conocimiento, dado que son compartidos por ambos lados del debate. Asimismo, esto garantiza que se trata de una auténtica controversia acerca del mismo fenómeno, lo cual evita que las críticas de una postura a otra se pasen de largo.

Las tesis anteriores no son controvertidas una vez que no consideramos como alternativa relevante al eliminativismo metodológico acerca del conocimiento. Es decir, una vez que se acepta el conocimiento como un *explanandum* relevante²⁵ para la epistemología y/o las ciencias cognitivas, la caracterización anterior se desprende con naturalidad; pero este presupuesto metodológico en sí mismo resulta controvertido. Por ello lo desarrollaré con un poco más de detalle en el siguiente apartado.

²⁴ El internista también podría extender las atribuciones de conocimiento a especies no-humanas formulando ciertos requisitos de manera adecuada, pero la cortesía se le extendería a muy pocas especies.

²⁵ Es decir, que las evaluaciones epistémicas hechas con la noción de conocimiento tienen un interés más allá de lo puramente sociológico. También cabe notar que la relación de Goldman con el naturalismo, acepta la tesis tradicional de que el conocimiento es fundamentalmente metalingüístico y por ende el concepto de conocimiento es el objeto a estudiar, pero a la vez considera que el contenido del concepto hace una referencia vaga a clases naturales.

1.3 El eliminativismo sobre el conocimiento

Una vez hecha la distinción entre el naturalismo sobre el conocimiento como postura metodológica y el naturalismo como postura metafísica es conveniente hacer algunas aclaraciones sobre el eliminativismo. Algo que es importante señalar es que aquí el eliminativismo sobre el conocimiento no es propiamente una postura metafísica, a pesar de que pudiera parecer lo contrario.

Esta impresión se debe a que el naturalismo epistémico²⁶, cuando se ocupa de los aspectos ontológicos, se compromete con que los objetos de estudio de la epistemología sean objetos, propiedades o procesos naturales.²⁷ Debido a este compromiso ontológico, el hecho de que la noción de conocimiento forme parte de la epistemología depende de que en el mundo natural exista algo que pueda identificarse —en algún sentido relevante— con el conocimiento.

No obstante, el eliminativismo sobre el conocimiento es en sí misma una tesis acerca de su utilidad teórica, como se puede ver en defensas del eliminativismo que o bien no son naturalistas (Bonjour, 2006) o son neutrales con respecto al naturalismo. Una postura similar es el deflacionismo sobre el conocimiento, donde éste simplemente se identifica con la creencia verdadera.²⁸ En el resto de este apartado expondré a grandes rasgos el eliminativismo metodológico, así como algunas razones para rechazarlo.

²⁶ Es decir, un naturalismo global adoptado por algún argumento o principio general. Uno también podría llegar a un naturalismo global examinando uno por uno los objetos de la epistemología y concluyendo que en cada caso son naturales, pero este no es el naturalismo global que tengo en mente aquí.

²⁷ No es trivial definir lo que se entiende por “natural” aquí, más adelante trataré este problema.

²⁸ Puede ser algo más complejo que la creencia verdadera, pero sigue siendo una noción mucho más débil en tanto que no incluye nada parecido a la justificación.

Hablaré en primer lugar del eliminativismo global de Stich, de manera breve. Considero que llevar a cabo este excursus vale la pena dado que el programa de Stich presenta problemas ineludibles, que no han sido respondidos enteramente y que es importante tomar en cuenta aquí. Intentaré justificar el rechazo al eliminativismo global sin dar una argumentación rigurosa. Posteriormente hablaré concretamente del eliminativismo acerca del conocimiento.

El eliminativismo global, como su nombre lo indica, es supresión de la epistemología (tradicional) como dominio de estudio. El pragmatismo propuesto por Stich (1994) conlleva un eliminativismo global donde se rechaza la idea de que pueda haber evaluaciones propiamente epistémicas —o de que éstas tendrían algún valor práctico en caso de que fuera posible hacerlas—en favor de evaluaciones utilitarias acerca de las prácticas y capacidades cognitivas. Para justificar este programa, Stich se apoya tanto en razones epistemológicas como metodológicas,²⁹ me concentraré aquí en las primeras.

De acuerdo con Stich, podemos evaluar cuáles son las mejores maneras de obtener creencias para un objetivo determinado (construir un puente, tener una sociedad más armónica, etc.), pero no cuál es el mejor estilo cognitivo *per se*, de manera que toda evaluación cognitiva es arbitraria. La tesis fundamental es que no se puede determinar objetivamente una meta epistémica valiosa, debido a que los candidatos principales para ello están subdeterminados y dicha subdeterminación sólo se podría suplementar de manera arbitraria. Un apoyo adicional consiste en mostrar cómo algunas nociones epistémicas usuales incluso son irrelevantes para el éxito práctico.

Stich ataca tanto la tesis de que las creencias verdaderas deben ser el objetivo de la epistemología, como la capacidad de que sea posible construir una noción adecuada de conocimiento o justificación aun suponiendo que las creencias verdaderas fueran una buena meta epistémica. El

²⁹ Las críticas metodológicas de Stich en el trabajo citado (1994) y otros posteriores sólo critica la epistemología como el estudio de conceptos epistémicos y no atañe al naturalismo epistemológico.

argumento de Stich en contra del valor de las creencias verdaderas es un argumento técnico partiendo de una teoría representacionista de las creencias (en concreto una teoría computacionalista al estilo de la tesis del lenguaje del pensamiento³⁰), donde éstas tienen algún tipo de existencia material que funciona como ejemplar. La valuación de una creencia para que ésta sea verdadera debe de ser determinada por medio de alguna interpretación sobre sus componentes (e.g. nombres). El candidato para dar una interpretación que ha sido defendido con más éxito en la literatura es una interpretación funcional/causal para determinar la referencia. Stich muestra entonces cómo puede construirse una función trivial "incorrecta" de cadenas causales que altera la referencia de un modo contraintuitivo. Pero de acuerdo con Stich no hay modo obvio más allá del sentido común de determinar una función de interpretación no arbitraria que agrupe a las cadenas causales adecuadas. Así, en palabras de Stich (1994, p. 101): "la mayoría de nosotros no habremos de encontrar ningún valor, ya sea intrínseco o instrumental, de tener creencias verdaderas." En resumen, el argumento de Stich busca socavar el valor que podría tener una creencia verdadera mostrando cómo se puede trivializar la noción.

No obstante, el conjunto de argumentos de Stich no establecen de manera concluyente que las creencias verdaderas no sean valiosas. En primer lugar, podría tomarse como un argumento en contra de determinadas teorías acerca de la creencia y de aquello que la interpreta como verdadera; por ejemplo, las teorías disposicionales o las teleosemánticas son alternativas que evitan este problema.³¹ Aun concediendo que estos problemas fueran insolubles, considero que habría que extraer una conclusión más modesta: que la creencia es una noción inadecuada para dar cuenta de una teoría normativa sobre las representaciones, quedando otras opciones abiertas tales como el

³⁰ Es decir, un funcionalismo donde el nivel relevante de descripción para explicar la intencionalidad es algún tipo de procesamiento simbólico: la tesis llamada por Fodor hipótesis del lenguaje del pensamiento.

³¹ Hay teorías disposicionales (e.g. Armstrong, 1973) que se concentran en la conducta que mostraría el sujeto de atribución de la creencia en un conjunto de escenarios relevantes, por lo cual no es necesario construir una función como la que describe Stich. Por otra parte las teorías teleosemánticas apelan a la selección natural, lo cual constriñe las cadenas causales y evita la subdeterminación. Por supuesto, estas teorías y otras que evitan los problemas del funcionalismo sobre creencias tienen sus propios problemas.

primitivismo sobre el conocimiento —donde éste no se define a partir de la noción de creencia— así como un programa naturalista que dispense de dicha noción sin abandonar una visión normativa y propiamente epistémica (es decir, no pragmática) acerca de las representaciones.³²

La existencia de representaciones correctas, sean o no caracterizadas como *creencias verdaderas*, es un primer paso necesario para recuperar nociones epistémicas más sustantivas como el conocimiento. No obstante, el paso de una cosa hacia la otra no es trivial. La noción de conocimiento va más allá de la mera corrección representacional: es necesario al menos que la creencia de que P (o algo análogo) haya sido adquirida y/o se mantenga de manera *epistémicamente* apropiada (o requisitos epistémicos que juegan un papel análogo), lo cual es justo el punto de contención. El eliminativista no acepta que haya una normatividad propiamente epistémica, es decir, un modo no-instrumental de clasificar a las creencias verdaderas de un sujeto en apropiadas y no apropiadas.

Recuperar las creencias verdaderas resuelve una parte importante del problema, ya que éstas se pueden identificar como la meta propiamente epistémica. Esto abre la puerta a posturas como el fiabilismo de proceso de Goldman

Me concentraré ahora en este tipo de eliminativismo. La primera alternativa es un eliminativismo en general acerca de las actitudes. El conocimiento proposicional —el punto de partida de esta investigación— se caracteriza como una actitud de un agente epistémico hacia una proposición, puesto que los enunciados tales como S sabe que P caen bajo la fórmula general de las actitudes, S ___ que P. Esto ha sido el marco de trabajo universalmente adoptado en la tradición del análisis

³² En todo caso sería más efectiva una crítica a la relevancia misma de la noción de representación tal como la de Rorty en *Philosophy and the Mirror of Nature* (1981). Para una presentación accesible sobre los problemas que conlleva dicha postura ver el intercambio entre Richard Rorty y Pascal Engel (2007).

del conocimiento proposicional. En una postura no-primitivista, lo que hace que el conocimiento sea una actitud es que uno de sus elementos es ser una creencia, mientras que en el caso del primitivismo el conocimiento es directamente un estado mental.

Esto se encuentra en cierta oposición al eliminativismo sobre las actitudes que sostiene que el vocabulario intencional utilizado en la psicología folk o vulgar (conocimiento, creencias, deseos) no tiene lugar en una descripción adecuada del fenómeno de la mentalidad (la cantidad de literatura sobre este punto es enorme). Bermudez (2006) en una crítica a Kornblith (2002) hace notar que un eliminativismo sobre las actitudes impide sostener la existencia del conocimiento como clase natural a partir de un argumento de un compromiso ontológico a partir de indispensabilidad teórica en las teorías sobre la cognición (la razón frecuentemente aducida es que ésta se explicará a un nivel menos abstracto, por medio de la neurociencia). Frente a esto cabe señalar dos cosas:

a) El argumento de indispensabilidad teórica no es el único modo de defender un naturalismo no-eliminativista. No resulta claro que sería inadecuado considerar como naturales cierto tipo de representaciones, ya sea como clase natural en una teoría no-esencialista, como bajo un criterio razonable de naturalidad.

b) El eliminativismo acerca de las actitudes suele partir de una cierta teoría representacionista, pero es compatible con otras teorías sobre las actitudes, tales como disposicionales o teleológicas.

Sobre el segundo punto, resulta evidente que no todas las disposiciones tienen que ser entidades o propiedades postuladas por una ciencia. En principio, parece que ciertas disposiciones extrínsecas son completamente ajenas a la ciencia. Por ejemplo, la disposición de ciertos alimentos altos en grasa de ser un éxito de ventas en los centros de comida rápida no tiene por qué ser una propiedad bien definida ni ser parte fundamental de ninguna ciencia. No obstante, existen otras disposiciones que resultan buenos candidatos para considerarse como naturales. Por ejemplo, el

hecho de que la sal se disuelva en agua (o un caso más problemático, que no se disuelva en benceno).

Sobre el eliminativismo general sobre actitudes, quisiera señalar un último punto. Es claro que un naturalista metafísico y/o metodológico podría seguir investigando el conocimiento (lo que vulgarmente se llama conocimiento) sin considerar que éste refiera de manera relevante a un fenómeno natural, del modo trivial como se pueden estudiar los fundamentos cognitivos de cualquier fenómeno humano. Esto sería similar a estudiar la agilidad de los monos araña, lo cual no implica considerar que se puede localizar a la *agilidad*, más que quizá de manera trivial (como un conjunto de disyunciones).³³ Para mostrar que el conocimiento es natural, no basta con mostrar que existen cierto tipo de representaciones que corresponden a lo que llamamos conocimiento, sino que dichas representaciones no son algo agrupado de manera artificial (como en el caso de la agilidad).

Finalmente quisiera examinar el caso del eliminativismo del conocimiento de acuerdo con el trabajo seminal de Quine (1969a), "Epistemology Naturalized." Para ello haré una breve reconstrucción de la postura de Quine, sin entrar con detalle a sus argumentos. Mi lectura es que Quine es un naturalista epistémico acerca del "conocimiento" científico³⁴, pero un eliminativista (o quizá un deflacionista) acerca del conocimiento ordinario. Más aún, este naturalismo tiene un carácter puramente normativo (como un canon de razonamiento) y no metafísico, y tampoco es un naturalismo metodológico para estudiar el conocimiento. Para distinguir le llamaré naturalismo-

c.

³³ Más aún, quizá ni siquiera sea posible encontrar condiciones necesarias y suficientes para una noción tan borrosa como la acrobacia.

³⁴ Más adelante se entenderá por qué pongo comillas de advertencia.

El naturalismo-c consiste en la idea de que las teorías exitosas³⁵ no requieren una justificación de su base empírica y sus métodos de generación de creencias fuera de la misma ciencia. Esto se puede ver en el ataque de Quine a la necesidad y posibilidad de una vindicación a priori de la verdad de ciertas creencias (para derivar de ahí el conocimiento científico). Una versión más fuerte de esta tesis, defendida por Colyvan (2010), es que estos modos de razonar son una buena manera de razonar per se. No obstante, el *conocimiento* como noción no aparece aquí como objeto de estudio; es decir, la pregunta acerca de qué es el conocimiento. Podríamos sustituir "conocimiento científico" por "las teorías científicas que sostenemos."³⁶ El naturalismo-c únicamente señala una autonomía epistémica de las ciencias, en tanto que no se requiere una justificación independiente a las teorías científicas de sus métodos y de su base empírica para sostenerlas. Esto es neutral con respecto al problema de la existencia y la naturalidad del conocimiento

Pero el naturalismo-c no es lo único que defiende Quine. Hacia el final de su artículo seminal, trata el tema de las evaluaciones epistémicas. Aquí es donde se puede ver claramente que Quine sostiene un eliminativismo de facto. El conocimiento en sí no juega ningún papel dentro de las evaluaciones epistémicas. Quine considera que el objeto de epistemología es examinar empíricamente la obtención de creencias y la eficacia de las distintas fuentes de obtención para que éstas sean verdaderas en un número satisfactorio, con el objetivo de lograr un mejoramiento epistémico. No obstante, el estudio del conocimiento incluye una explicitación de cuáles son estos criterios satisfactorios; pues notemos que un método puede otorgar una mejor tasa de creencias verdaderas, pero un número mayor de creencias falsas.³⁷ ¿Qué criterios deben cumplirse para que una creencia verdadera sea conocimiento o siquiera justificación? Si pensamos que hay una respuesta a esta

³⁵ O en un holismo como el de Quine, el conjunto de la ciencia.

³⁶ Podría agregarse que sostenemos como verdaderas o al menos como verosímiles (al estilo de Popper o Hilpinen), pero en realidad basta con considerarlas razonables para sostenerlas.

³⁷ Por ejemplo, un método A podría proveer al agente con 9 creencias verdaderas y 1 falsa (90% de éxito), mientras que el método B proveería 15 creencias verdaderas y 5 falsas (75% de éxito). Puede haber razones pragmáticas para preferir uno de los dos, pero decidir cuál es en sí mismo epistémicamente mejor no es un asunto trivial.

pregunta, igualmente sería necesario dar una respuesta acerca de cómo se individualizan dichos métodos, tal como se ha tratado en las teorías fiabilistas. No obstante, en el texto citado Quine no busca una respuesta a esta pregunta porque su naturalismo epistemológico es una propuesta metodológica acerca de creencias verdaderas y de evaluar el éxito de distintas maneras de obtenerlas, no de elucidar qué es conocimiento o siquiera la justificación.

En resumen, puesto que dejan el dominio epistémico sin una explicación satisfactoria, considero que existen buenas razones para rechazar el eliminativismo de Stich global y el deflacionismo de Quine. En el caso de Stich, sus argumentos dependen de presupuestos teóricos que no es obligatorio adoptar, y en el caso de Quine no hay realmente un argumento para dejar de lado el estudio de las condiciones de satisfacción para que haya conocimiento, sino únicamente un rechazo a la idea de que haya que fundar el conocimiento científico a priori.

Prima facie el conocimiento juega un papel primordial en las prácticas epistémicas por medio de evaluaciones realizadas en las prácticas epistémicas cotidianas³⁸ y en la epistemología (como parte de una larga tradición histórica), y su papel comienza a crecer en disciplinas como la psicología (J. Nagel, 2013), además de que su uso es fundamental en la etología cognitiva (Kornblith, 2002).³⁹ Con relación al eliminativismo sobre las actitudes, quizá una teoría final sobre la cognición no incluirá un vocabulario intencional o actitudinal, pero hay buenas razones para dejar una discusión sobre este punto fuera del enfoque principal de este trabajo. El uso de la noción de conocimiento en la psicología y la etología es un punto de partida apropiado para investigar la naturalización del conocimiento desde el punto de vista de la filosofía de las ciencias cognitivas, sin negar que el valor de otros trabajos especulativos sobre cómo será una teoría final de la cognición.

³⁸ Es decir, hablo aquí de la noción "folk" de conocimiento, si bien queda por demostrar si ésta es adecuada.

³⁹ En rigor, sólo se puede afirmar categóricamente que el término "conocimiento" aparece en estas disciplinas. Hace falta discutir si realmente es la misma noción, pero al menos no es evidente que no lo sea.

1.4. Naturalidad

En este apartado examinaré la tesis de naturalidad del conocimiento (CN). ¿Qué implica decir que un término epistémico como el conocimiento es natural? La manera más fuerte es identificar al *conocimiento* reductivamente con una clase natural, una manera menos fuerte es mostrar cómo el conocimiento se fundamenta parcialmente en determinadas clases naturales, tal como sucede en algunas propuestas de epistemología de la virtud naturalizada.

Aquí considero necesario responder a dos preguntas, la primera de ellas es ¿por qué recurrir a este tipo de nociones, como la identificación reductiva o la fundamentación? La idea es que dichas clases naturales jueguen un papel *constitutivo* para el conocimiento. No basta con señalar que el conocimiento pueda explicarse o estudiarse empíricamente con determinadas ciencias. Por ejemplo, puede haber una explicación de ciertos aspectos de lo que hacen los matemáticos desde el punto de vista de sus procesos psicológicos, pero esto por sí mismo no apoya un psicologismo acerca de la naturaleza de la demostración matemática. Teniendo en mente este objetivo de explicar el papel constitutivo que tienen las clases naturales en el conocimiento en una visión naturalista, será necesario dar una caracterización más precisa de lo que se entiende por clase natural.

La segunda pregunta, ¿por qué no simplemente tomar como clase natural los términos utilizados por determinadas ciencias empíricas? Esto resulta inicialmente plausible, pero en última instancia es insatisfactorio si dichos términos no son indicativos de algo más. El hecho de que una teoría psicológica utilice la noción de conocimiento podría tomarse como evidencia de que el conocimiento es una clase natural, pero también se puede hacer una interpretación en la dirección contraria: que dicha teoría es inadecuada por utilizar una noción epistémica que no es natural. Por

ello no basta con apelar a los términos teóricos de las ciencias empíricas, sino a tener un determinado compromiso sobre la referencia de dichos términos.

Antes de entrar con mayor detalle en el mejor modo de caracterizar lo que es una clase natural, hablaré de un requisito previo: la tesis de que el conocimiento —sea natural o artificial— es localizable. Es decir, tiene una descripción equivalente en un lenguaje fundamental (presumiblemente, un lenguaje fisicalista). Como se verá a continuación, el proyecto de localización es demasiado débil como para considerarlo una naturalización efectiva del conocimiento, pues dicha localización incluye también a las clases artificiales. No obstante, como explicaré posteriormente, la localización puede ser parte de la estrategia de naturalización (tal como sucede en la propuesta de Kornblith). Además de señalar que es insuficiente para caracterizar una postura naturalista, mostraré que es un presupuesto razonable para adoptar como parte de un marco neutral para plantear la discusión entre naturalismo y no-naturalismo.

Lo que podría llamarse una tesis de localizabilidad global⁴⁰, tal como aparece en Jackson (2000)⁴¹ y otros autores afines, consiste en realizar un trabajo de organización conceptual que toma en cuenta a las ciencias empíricas con el fin de evitar el eliminativismo sobre la mayor cantidad de entidades posibles y al mismo tiempo mantener una ontología razonablemente austera. La idea es que, si bien es razonable tomar algunas cosas como fundamentales, tales como el espacio-tiempo y algunas partículas físicas (o si se prefiere, campos⁴²)— sería poco razonable tener una ontología donde todo fuera fundamental.

⁴⁰ Como estrategia filosófica, el proyecto de localización de Jackson es completamente diferente a una estrategia naturalista como la que propone Kornblith

⁴¹ Aquí no me refiero al proyecto de localización de Jackson sobre conceptos folk, que en cierta manera es contrario a la estrategia de naturalización que explicaré posteriormente.

⁴² Ver Hobson (2004) "There are no particles, only fields" para una defensa de que la teoría cuántica de campos presenta una ontología más razonable que la mecánica cuántica a pesar de ser empíricamente equivalente.

Partiendo de la tesis plausible de que algunas entidades y propiedades físicas son los candidatos adecuados para ser consideradas como fundamentales, la idea de la localización es la siguiente: intentar mostrar cómo las entidades y propiedades no-fundamentales se pueden derivar de una base ontológica fundamental utilizando herramientas lógicas y matemáticas, particularmente funciones. Para capturar esta intuición, Jackson utiliza la noción de superveniencia global. En el caso del conocimiento, también resulta relevante la noción independiente de superveniencia epistémica.⁴³ Se trata de dos tesis distintas. La primera intenta capturar la dependencia de toda la ontología sobre la física básica, mientras que la segunda —una consecuencia lógica de la primera, pero que puede defenderse de manera independiente— busca capturar la dependencia de lo epistémico sobre lo no-epistémico.

No resulta trivial formular una tesis satisfactoria de superveniencia global que capture esta relación de dependencia con una noción puramente intensional, es decir definida con una semántica de mundos posibles. La idea es que las variaciones entre mundos posibles ilustran, al menos en parte, esta dependencia. Debido a las limitaciones de la semántica de mundos posibles existen propuestas de capturar esta relación de *fundamentación* [grounding] de un modo hiper-intensional (Schaffer, 2009). No obstante, como hace notar Jackson, para el proyecto de localización no es indispensable formular una tesis que capture la dependencia con precisión.⁴⁴ Jackson propone una tesis más débil que no captura por completo la idea de superveniencia, pero que es una consecuencia de la superveniencia global y que además sería una tesis que no aceptaría alguien que no acepte el materialismo. La tesis de Jackson, que llamaré dependencia global de lo físico (DGF), y que es consecuencia de la superveniencia global de lo físico (SGF) es la siguiente:

⁴³ Para formulaciones de la tesis de superveniencia epistémica ver Van Cleve (1999), Kim (1988), y Turri (2010).

⁴⁴ Se trata de un objetivo importante para la metafísica, pero el proyecto de localización es más modesto.

DGF: Cualquier mundo que sea un duplicado físico de nuestro mundo, es un duplicado *simpliciter* de nuestro mundo.

Como se puede ver, DGF rechaza que en el mundo actual exista una especie de superávit de lo físico (por ejemplo, rechaza un dualismo de propiedades con respecto a lo mental). Una tesis de dependencia epistémica de lo físico sería la siguiente:

DEF: Cualquier mundo que sea un duplicado físico de nuestro mundo, es un duplicado epistémico de nuestro mundo.

A continuación, daré una justificación para adoptar DEF para investigar la metafísica del conocimiento. Comenzaré por responder a algunas desventajas. La primera es que, al igual que DGF, DEF es más débil que la relación de superveniencia, pues no captura las relaciones de necesidad entre ambos conjuntos de propiedades a través de contrafácticos. No obstante, con DEM sólo se pretende descartar que las propiedades epistémicas sean un tipo de propiedades fundamentales en el mundo actual.

La segunda desventaja es que DEF en otro aspecto es más fuerte de lo que se requiere. Puesto que el conocimiento tiene un componente representacional, las propiedades epistémicas dependen, al menos en parte, de las propiedades mentales. Por ello, DEM tiene la consecuencia de que al menos estas propiedades mentales, tienen una dependencia de lo físico, de modo que DEM está tomando postura acerca de la naturaleza de lo mental, lo cual es irrelevante para la discusión sobre la naturalidad del conocimiento en tanto que se opone a la artificialidad del mismo. Considero que esto es inocuo por dos razones. En primer lugar, considero que la controversia más relevante acerca de la reducción de las propiedades mentales a lo físico es con relación a los estados cualitativos, y que la creencia (y en general el tipo de representación que requiere el conocimiento

proposicional) es menos problemático. Pero aun concediendo que algunos aspectos intencionales o representacionales de la mente no pudieran caracterizarse de manera física —como sostiene, por ejemplo, Haugeland (1997)⁴⁵— esto no afectará mayormente a la presente investigación, puesto que la idea principal es que lo epistémico no es fundamental. Si hubiera que agregar algunas leyes psicodinámicas al conjunto de entidades y propiedades fundamentales, simplemente habría que sustituir DEM por una tesis sobre la dependencia de lo epistémico en lo material con adición de estas entidades y propiedades.

Una vez resueltas estas desventajas, sólo hace falta dar cuenta de la neutralidad de DEM para plantear el problema de la naturalidad del conocimiento. Para ello basta con hacer notar que tanto lo natural como lo artificial se realiza de un modo material (esto es una tesis más débil que señalar que se reducen a lo material), lo cual incluye tanto entidades, propiedades e incluso leyes (como hace notar Fodor acerca de las ciencias especiales). Me parece que basta con un ejemplo de Kim (1992) para mostrar esto⁴⁶. Históricamente, diversas piedras fueron clasificadas como *jade* por su apariencia. No obstante, con la existencia de herramientas más precisas se pudo notar que las piedras clasificadas como jade pertenecían a uno de dos compuestos químicos: jadeíta y nefrita. El jade sigue existiendo como clasificación, pero corresponde a razones históricamente contingentes (e.g. que hubiera acceso a ambos compuestos y no sólo a uno de ellos) y a intereses prácticos (desde el punto de vista estético son equivalentes). Por otro lado, la clasificación en jadeíta y nefrita obedece a criterios objetivos: todo el jade tiene la composición $\text{NaAlSi}_2\text{O}_6$, mientras que la nefrita tiene una composición $\text{Ca}_2(\text{Mg}, \text{Fe})_5\text{Si}_8\text{O}_{22}(\text{OH})_2$. Pero también el jade, a pesar de ser una clasificación artificial, se realiza físicamente, ya sea como $\text{NaAlSi}_2\text{O}_6$ o como $\text{Ca}_2(\text{Mg}, \text{Fe})_5\text{Si}_8\text{O}_{22}(\text{OH})_2$. Aquí se puede ver como la tesis de dependencia global

⁴⁵ Específicamente, sobre la *comprensión*, un término que no aparece en el estudio empírico y filosófico sobre la cognición de las tradiciones más importantes.

⁴⁶ Aunque la perspectiva de Kim (2005) acerca de este problema no es metafísica, sino epistémica (sobre la proyectabilidad de los términos científicos para hacer inferencias).

ilustra esto, ya que si todo —incluyendo lo químico— depende ontológicamente de lo físico fundamental,⁴⁷ en todo mundo que fuera un duplicado físico del nuestro, los enunciados "x es jadeíta", "x es nefrita" y "x es jade" serían verdaderos o falsos para exactamente todos los duplicados de cada objeto del mundo actual. El ejemplo anterior ilustra la discusión real de ciencias cognitivas, por ejemplo, sobre si la clasificación de las emociones (enojo, tristeza, etc...) corresponde a procesos cerebrales; donde la evidencia hasta ahora indica que no. Es decir, que la clasificación usual de las emociones es hasta cierto punto arbitraria (Griffiths, 2008).⁴⁸

Ahora bien, del hecho de que sea posible en principio localizar una clase artificial, no se sigue que sea algo práctico o que tenga un interés teórico. En el caso del jade, se puede caracterizar por medio de una disyunción entre dos compuestos químicos, pero en otras clases artificiales puede haber relaciones funcionales muy complejas, incluyendo elementos contextuales, de modo que la localización sería imposible en términos prácticos. Por ello, a pesar de señalar que el conocimiento es localizable en principio (sea natural o artificial), queda como una cuestión abierta que dicha localización sea factible.

Una vez establecido que el conocimiento es localizable, y habiendo señalado con un ejemplo la distinción intuitiva entre lo natural y lo artificial, es necesario indicar de qué manera puede entenderse la tesis de la naturalidad del conocimiento. Para ello, sugiero considerar las siguientes alternativas:

⁴⁷ Con ello no señalo que la química como ciencia se reduzca a la física, pues esto no sería una tesis de dependencia meramente ontológica, sino nomológica.

⁴⁸ Si bien sería posible rescatar la idea de que la naturalidad de las emociones si la psicología fuera nomológicamente independiente de las neurociencias (como sostiene Fodor) y una clasificación de emociones similar a la *folk* tuviera un papel en ella.

a) El conocimiento se identifica reductivamente con una clase natural o una disyunción manejable⁴⁹ de clases naturales (CN1).

b) En la definición de conocimiento aparecen de manera explícita o implícita un conjunto de clases naturales (CN2).

La primera tesis es la más fuerte. La afirmación de que el conocimiento es una clase natural, tomada de modo llano, es prácticamente vacía; difícilmente podría haber una discusión en torno a la naturalidad del conocimiento sin una propuesta más concreta. El modo más fructífero de entenderla dicha afirmación por medio de una identificación reductiva con una propiedad biológica, psicológica, etc., la cual a su vez es una clase natural. Sin entrar de lleno en las distintas teorías sobre cómo caracterizar una identificación reductiva, la noción indica una relación de identidad asimétrica, en la que uno de los términos es más básico o más fundamental que el otro. Así, si bien el H₂O es idéntico al agua, no decimos que el H₂O se reduce al agua, sino únicamente lo converso. Dicha prioridad puede entenderse en términos de *explicación* (Garfinkel, 1981). Mientras que el H₂O explica las propiedades macroscópicas del agua de manera microestructural, lo converso no se cumple. Si el conocimiento se identifica con una sola clase natural, será lo que podríamos llamar una naturalización perfecta. Si se descubre que el término de "conocimiento" en realidad tiene como referencia un pequeño número de clases naturales, se tratará de una naturalización imperfecta (por ejemplo, si descubriéramos que el conocimiento perceptual y el conocimiento inferencial son dos clases naturales distintas). Lo importante es que dicha disyunción de clases naturales no consiste simplemente en mostrar los realizadores del conocimiento, sino que verdaderamente exista una identificación reductiva con cada uno de ellos. (Por supuesto, si la disyunción comenzara a crecer desmedidamente deberíamos empezar a sospechar que algo salió mal).

⁴⁹ Es decir, lo contrario a lo que Fodor (1974) llama una "disyunción salvaje."

A continuación, será necesario revisar brevemente distintas posturas sobre cuál es la manera adecuada de formular CN1. No hay que presuponer que estas formulaciones compiten por ser el modo universal de caracterizar la noción de clase natural. Un enfoque pluralista resulta más fructífero: podemos utilizar distintas nociones de clase natural para distintos fines, lo cual es acorde con el distinto carácter que tienen los objetos de distintas ciencias (e.g. las propiedades biológicas son mucho más variables que las propiedades físicas).

Los criterios satisfactorios que debe cumplir una teoría de clases naturales son fundamentalmente dos. El primero de ellos es metafísico: consiste en indicar cuáles son las agrupaciones de objetos, propiedades y procesos que son hasta cierto punto objetivas⁵⁰, o en las posiciones realistas —como indica una metáfora usual— que recortan al mundo por las junturas. El segundo criterio es epistémico: las clases naturales son proyectables, permiten realizar inferencias inductivas y formar parte de explicaciones científicas. Una tercera función de la noción de clase natural es semántica, como sucede en las teorías externistas sobre la referencia. Me concentraré aquí de las primeras dos. Ambas son tomadas de trabajos que defienden una postura explícitamente realista sobre clases naturales, pues a partir de un nominalismo de clases naturales⁵¹ la discusión metafísica sobre la naturalización del conocimiento sería trivial. Las siguientes son caracterizaciones de qué es una clase natural:

a) Condiciones necesarias y suficientes (i.e. esencialismo)⁵²

⁵⁰ También en las posturas anti-realistas hay cierta objetividad en las clases naturales. Aún si las clases naturales no capturan divisiones que existen realmente en la naturaleza, no son un producto de condiciones históricas contingentes o de intereses extra-científicos.

⁵¹ Para dos ejemplos clásicos, ver Quine (1969b), "Natural Kinds" o Goodman (1954), *Fact, Fiction, and Forecast*.

⁵² Boyd considera que su teoría de cúmulos no sólo explica las clases proyectables utilizadas por la ciencia, sino que éstas son esencias metafísicas reales (es decir, no son meramente nominales). Sin embargo, como suele hacerse, utilizaré el término "esencialismo" para referirme al primer grupo de teorías.

b) Cúmulos de propiedades causales homeostáticas

El esencialismo de clases naturales se puede formular en términos de necesidad y suficiencia. De acuerdo con una tradición iniciada por Kripke (1980), los términos de clases naturales —al igual que los nombres— se caracterizan por designar rígidamente, es decir, por designar lo mismo en todos los mundos posibles. La esencia de una clase natural consiste en el conjunto de propiedades que los miembros de dicha clase poseen necesariamente, es decir, en todos los mundos. La suficiencia la provee el aspecto constitutivo que tiene la esencia: "los electrones necesariamente son partículas con carga... si una partícula carece de este poder causal, sin importar lo que sea —de acuerdo con el esencialismo— no sería un electrón" (Ellis, 2007, p. 13). Una tesis comúnmente defendida es que la esencia de las clases naturales corresponde a propiedades microestructurales, por lo cual las clases artificiales, aún si designan rígidamente —como sostienen LaPorte (2004) y Kornblith (1980)— no podrían tener una esencia más que nominal.

El problema con la visión esencialista de las clases naturales para la presente investigación radica en la dificultad de incluir a las clases biológicas y psicológicas, que son las relevantes para discutir la naturalización del conocimiento. En el caso de muchas clases biológicas, como las especies⁵³, existe un elemento histórico que no queda capturado por las propiedades microestructurales. Si en otro mundo posible los mamíferos evolucionaron hasta ser iguales a los insectos de nuestro mundo, no serían insectos (LaPorte, 2004), al menos si tomamos a las especies desde un punto de vista cladístico.⁵⁴ Aún en el caso de otro tipo de clases biológicas, como las rutas metabólicas, la reducción completa a propiedades microestructurales resulta controvertida. También se ha

⁵³ Hay otra visión sobre las especies biológicas que no considero aquí, según la cual son individuos extendidos espaciotemporalmente, compuestos de otros individuos.

⁵⁴ También la función biológica presenta problemas similares. Si caracterizamos a la función biológica como la función propia propuesta por Millikan (1995) surgen los mismos problemas que en el caso de las especies. Si utilizamos el análisis funcional, habría que justificar por qué éste constituye una clase natural, ya que en éste un sistema es conceptualizado según intereses explicativos.

considerado que en el caso de las especies el cambio a través del tiempo introduce un problema de vaguedad incompatible con el esencialismo sobre clases naturales (para una opinión en contra, ver Sober, 1980). En el caso de las clases psicológicas, sería necesario justificar una serie de reducciones interteóricas entre clases,⁵⁵ siendo una de las más problemáticas la identificación entre tipos psicológicos y otros más fundamentales (como neurológicos y/o biológicos). Asimismo, como hace notar Sober (1980), el requerimiento del esencialismo es definir una especie a partir de las propiedades compartidas por los organismos miembros, no obstante la biología define a las especies por medio de conceptos poblacionales como el de acervo genético, donde la población a su vez no se define por medio de los organismos que la componen.

Por razones como éstas, algunos teóricos de clases naturales como esencias optan por considerar únicamente como clases naturales a los componentes físicos más básicos (e.g. Ellis, 2007) o bien una noción más amplia de clase natural que incluye a clases menos fundamentales (como los elementos químicos), pero sin incluir a las clases biológicas. Por ello la versión esencialista como hace notar Kornblith (2002) no es adecuada para caracterizar CN (especialmente si la naturalización apela a nociones evolutivas).

La segunda alternativa para caracterizar clases naturales procede de una propuesta de Boyd, como cúmulos causales homeostáticos. Esta noción de clases naturales es muy relevante para la discusión no sólo debido a que admite clases biológicas y psicológicas con mayor facilidad que el esencialismo; sino porque es una noción explícitamente adoptada por defensores de CN1 y CN2, como Kornblith y Millikan. Ésta es la caracterización de cúmulo de propiedades homeostáticas que hace Boyd en "Homeostasis Species and Higher Taxa" (1999, p. 142),

⁵⁵ No presupongo que exista una jerarquía perfectamente definida de niveles ontológicos y/o nomológicos al estilo de Oppenheim y Putnam (1958). Es decir, considero que el agua es una clase más fundamental que las mitocondrias, pero esto no implica agrupar a todo lo químico y a lo biológico en niveles. A mi modo de ver, las reducciones interteóricas se dan clase por clase.

- a. Existe una familia de propiedades F que están contingentemente agrupadas, forman un cúmulo en el sentido de que co-ocurren en un número importante de casos.
 - b. La co-ocurrencia es típicamente el resultado de
 - i) o bien una homeostasis metafórica (a veces literal), en el sentido de que las propiedades en F tienden a favorecer la presencia unas de otras;
 - ii) o bien mecanismos y/o procesos que mantienen la presencia de las propiedades F,
 - iii) o ambas.

Además, las siguientes características (pp. 142-143) resultan en lo que a mi modo de ver es por una parte, una definición estipulativa de clase natural; y por otra, muestra cómo los cúmulos pueden tener la función que tradicionalmente tienen las clases naturales:

- c. La agrupación de propiedades F es causalmente importante, tienen efectos teóricos y prácticos producidas por la ocurrencia conjunta de las propiedades F junto con todos o algunos de los mecanismos que las producen.
- d. Existe un término de clase *t* que se aplica a las cosas en las que el cúmulo de propiedades F ocurre.
- e. *t* no se puede definir analíticamente, sino sólo mediante la totalidad o parte del cúmulo F y algunos o todos los mecanismos subyacentes
- f. La importancia causal del cúmulo F junto con los mecanismos subyacentes hacen que la propiedad denotada por *t* sea una clase natural. Ningún refinamiento en el uso lingüístico que reemplace a *t* por un término menos vago (desde el punto de vista extensional) preservará la naturalidad de la clase descrita por *t*; en el sentido de que dicho refinamiento le dará importancias a distinciones que son irrelevantes para la explicación causal o la inducción.

Existen otras consecuencias lingüísticas y epistémicas que hace notar Boyd en las que no hace falta profundizar aquí, tales como la indeterminación extensional (la existencia de casos en los que no se puede establecer si un individuo pertenece a F), el hecho de que las propiedades en F se conozcan a posteriori, etc. Lo importante que cabe notar aquí es que los cúmulos de Boyd permiten el tipo de clases adecuado para naturalizar el conocimiento, y que dicha naturalización no sea trivial desde el punto de vista metafísico.

1.5. Estrategias de naturalización

A continuación mostraré algunas estrategias de naturalización para las dos tesis CN1 y CN2. Primero señalaré cuál es la estrategia del proyecto de localización para mostrar por qué es inadecuada para naturalizar el conocimiento. Anteriormente ya establecí por qué la localizabilidad es más débil que la tesis de que el conocimiento es natural, pero además quiero mostrar cómo una metodología de localización del concepto preteórico de conocimiento es contraria a los objetivos de CN1. Posteriormente reconstruiré lo que considero que es la estrategia de naturalización tomada de Kornblith (2002), señalando algunos de sus problemas. Finalmente propondré una estrategia de naturalización para CN2.

De forma simplificada, la estrategia de localización, tal como la propone Jackson (2000) y Braddon-Mitchell (2008) para la localización de una noción t adaptado para el caso de los términos epistémicos es la siguiente:

1. Ramificar t con predicados que no contengan términos epistémicos.
2. Encontrar los realizadores de los predicados de la ramificación de t a partir de un conjunto de entidades fundamentales.
3. Si se encuentran suficientes realizadores, concluir que t ha sido localizado.
4. Si no hay suficientes realizadores, concluir que t no existe.

El problema de dicha estrategia, además de que la realización al no ser una identificación reductiva no constituye un criterio suficiente de naturalización (como expuse anteriormente) es que le da un papel excesivamente importante a la noción preteórica. La naturalización, por el contrario, no busca conservar a toda costa todos los elementos de una noción preteórica y además está abierta a que la clase con la cual se realice la identificación tenga características nuevas que puedan

descubrirse a posteriori.⁵⁶ Contrastemos esto con la estrategia que sigue Kornblith (2002), que podríamos llamar de fijación de referente:

1. Considerar las características descriptivas y normativas de t junto con sus capacidades explicativas y evaluativas.
2. Utilizar la noción inicial de t para fijar un referente C .
3. Mostrar que C es una clase natural.
4. Mostrar que C puede realizar las funciones explicativas y evaluativas de t .
5. Afirmar la existencia de una identificación reductiva.

Esta última estrategia, a diferencia de la estrategia de localización, sí cumple con los criterios para realizar una naturalización efectiva. Aquí no se apela a una noción preteórica, sino a una noción inicial, que puede estar en menor o mayor grado cargada de teoría. También puede ser que t no tenga ninguna carga teórica relevante, pero esto no le da un estatuto especial a t . Es decir, lo que se trata es de emular la práctica científica. Si queremos saber con qué características naturales identificar al oro, comenzaremos con lo que sepamos acerca del oro, pueden ser nociones preteóricas como su color, pero también el hecho de que se disuelve en aqua regia. Esta noción inicial fija un referente, al modo de las teorías externistas semánticas. Una vez que se fija dicho referente, importan poco las características de la noción inicial. La noción preteórica juega un papel mínimo. No es necesario capturar ninguna de las características iniciales de t ; por ejemplo, podríamos partir de que el conocimiento está relacionado con la reflexión, pero una vez hecha una identificación reductiva, podríamos descubrir a posteriori que el conocimiento realmente no incluye a la reflexión. No obstante, resulta importante mostrar que C puede tener funciones normativas y explicativas similares a las de t , mostrando así que se conserva la noción epistémica t

Si bien esta estrategia funciona en términos generales, considero que es posible mejorarla. El problema con la estrategia de Kornblith es que sólo se preocupa por mostrar que C cumple con

⁵⁶ Estoy presentando una versión de la estrategia de localización que es sumamente liberal permitiendo la realizabilidad múltiple, y que por tanto se inclina hacia mantener los términos preteóricos. No obstante, muchos de los defensores de esta estrategia (en ocasiones los mismos Braddon-Mitchell y Jackson) la describen como algo más cercano a la estrategia de acomodación.

las funciones de t de una manera aproximada, particularmente para rebatir la afirmación de que una noción naturalizada de conocimiento no puede tener un papel normativo. Considero que es necesario un cierto trabajo de acomodación, no a partir de una noción preteórica ni tampoco de una noción inicial de t , sino de un trabajo más preciso donde se establezcan las condiciones de adecuación de t tal como sucede en los análisis epistemológicos. Agregó además una propuesta de carácter metafilosófico: estas condiciones de adecuación de t no son simplemente un desarrollo del contenido preteórico de t , sino que t dentro de una teoría epistemológica es un término estipulado con el fin de realizar de la manera más adecuada posible el trabajo teórico que t preteórico realiza imperfectamente (sin importar qué tanto contenido compartan).

Estas condiciones de adecuación no están escritas sobre piedra, pues de otro modo sería simplemente localización, pero se deben reescribir con más cuidado del que indicaría una estrategia como la de Kornblith. Por ello propongo tanto la comparación de t con las clases candidatas para realizar una identificación reductiva, considerando distintas teorías epistemológicas sobre t . Por ejemplo, el primitivismo sobre el conocimiento permitiría identificarlo con determinadas clases naturales, pero si el conocimiento es una creencia con una determinada propiedad las clases candidatas serían otras. De acuerdo con lo que digo, habría que evaluar ambas alternativas (junto con otros elementos) para llegar al mejor candidato para una identificación reductiva, y posteriormente juzgar a dicho candidato. La estrategia que propongo para naturalizar el conocimiento bajo la tesis CN1 quedaría del siguiente modo:

1. Considerar las características descriptivas y normativas de t junto con sus capacidades explicativas y evaluativas.
2. Enlistar las clases naturales C_1 , C_2 , etc que sean buenos candidatos para identificarse reductivamente con el conocimiento.
3. Tratar en la medida de lo posible de escoger los candidatos que puedan cumplir con los desiderata de 1 *desde el punto de una o varias teorías epistemológicas que provean las condiciones de adecuación de t .*
4. La noción epistemológica de t en 3 sólo tiene una función de guía, será posible revisarla mientras se rescaten los elementos principales en 1.
5. Juzgar si los pasos 3 y 4 son suficientemente satisfactorios como para considerar que existe una identificación reductiva, así como señalar algunas de sus consecuencias.

Me parece que esta estrategia es un buen punto medio entre las dos descritas y que no sesga el debate en contra del naturalismo. La idea de que la práctica científica consiste simplemente en estudiar los referentes que fijan nuestras nociones antecedentes es una ficción. Este mismo proceso de acomodación sucede también en la ciencia. Fue posible identificar al oro con el número atómico 79, porque ya existían condiciones de adecuación previas (relacionadas con sus interacciones con otros elementos y compuestos, su punto de fusión, etc) que permitieron seleccionar a los ejemplares adecuados para probar esta hipótesis reductiva. Por ejemplo, si no supiéramos que el hielo y el agua líquida se transforman mutuamente (una de las condiciones de adecuación para identificar algo como agua), analizaríamos únicamente ejemplares de agua en estado líquido, las cuales compartirían una serie de propiedades que el agua no posee en estado sólido. Si Kripke está en lo correcto, una noción como "este metal amarillo" puede ser suficiente para fijar un referente, pero de ahí no se sigue que sabremos lo suficiente sobre dicho referente para seleccionar los ejemplares adecuados para estudiarlos. Del mismo modo, los análisis de la epistemología tradicional son condiciones de adecuación, no como elucidaciones a priori de nociones preteóricas, sino que están informados por lo que sabemos sobre los mecanismos y las prácticas epistémicas.

1.6. Conclusión del capítulo

La pregunta ontológica por la naturalidad del conocimiento es un problema distinto, si bien no completamente independiente, del naturalismo metodológico. Una vez aclarado esto, ha sido necesario dar algunos desiderata generales sobre el fenómeno a estudiar, el conocimiento proposicional, así como precisar la tesis de la naturalidad del conocimiento. He intentado ser lo más austero posible al dar estos desiderata, para tener una noción sin compromisos con alguna

teoría epistemológica particular y que no se encuentre demasiado delimitada por nuestras intuiciones preteóricas sobre qué es el conocimiento.

Dado que este trabajo buscará defender un naturalismo no-eliminativista, uno de los objetivos de este capítulo ha sido motivar esta postura. De acuerdo con el naturalismo no-eliminativista, la noción de conocimiento no es una parte de nuestro bagaje precientífico que deba desaparecer en un estudio empírico sobre las representaciones adecuadas, como parecen proponer algunas propuestas sobre naturalismo metodológico. Si esta tesis es correcta, el conocimiento —junto con sus funciones evaluativas y normativas— debe ser parte de una descripción científicamente adecuada del mundo natural.

Asimismo, he buscado precisar un modo de entender la tesis de naturalidad del conocimiento. Como señalé anteriormente, el esencialismo sobre clases naturales resulta inadecuado en el ámbito de la ciencia cognitiva, y quizá de manera más amplia, en lo biológico. Para ello, al igual que Kornblith (2002) he retomado la teoría de cúmulos de Boyd. Asimismo, añado la posibilidad de hablar de grados de naturalización de Lewis (1983), donde la naturalidad de una clase se define a partir de su cercanía con las clases perfectamente naturales que toma como primitivos. Posteriormente, señalé algunos elementos acerca de lo que implica naturalizar un término, mostrando que la identificación reductiva presente en la literatura no es la única alternativa viable.

Finalmente, di una clasificación a grandes rasgos sobre las estrategias de naturalización que considero más importantes para la discusión que llevaré a cabo en los siguientes capítulos: la de localización, la de identificación de referentes y la de acomodación. Señalé también algunas razones por las cuales la estrategia de acomodación es la más adecuada para conservar el fenómeno a estudiar, pero sin que los elementos preteóricos tengan un papel inadecuadamente grande en la investigación sobre la naturalidad del conocimiento.

CAPÍTULO 2: FIABILISMO ECOLÓGICO

2.1. Introducción

El objetivo de este capítulo es reconstruir y evaluar críticamente la propuesta de Hilary Kornblith sobre cómo naturalizar el conocimiento. Kornblith ha defendido en diversos trabajos (1999a, 1999b, 2002, 2005, 2006) un proyecto que tiene una parte metodológica —el conocimiento se debe estudiar empíricamente— y una parte ontológica —el conocimiento es una clase natural (CN). La tesis que me interesa para fines de esta trabajo es la segunda, pero dado que en el trabajo de Kornblith ambas tesis se apoyan mutuamente, será necesario hacer referencia en ocasiones a la tesis metodológica.

En *Knowledge and its Place in Nature* (2002), Kornblith realiza la propuesta más completa y sistemática para naturalizar el conocimiento proposicional que se pueda encontrar en la literatura de las últimas décadas. Señalaré a continuación cuál es la propuesta a grandes rasgos, haciendo notar cómo cubre los aspectos más importantes de un proyecto de naturalización. En este trabajo y los otros citados, Kornblith provee un marco meta-filosófico, donde sostiene que el objeto de estudio de la filosofía —y en concreto, de la epistemología— no son nuestros conceptos (i.e. el concepto de conocimiento), sino el referente mismo de dichos conceptos (i.e. el conocimiento). Esta meta-filosofía incluye una defensa de una metodología filosófica naturalista, mostrando cómo es posible superar la epistemología centrada en conceptos. Kornblith considera que, partiendo de una semántica externista, nuestros conceptos fijan un referente (esto lo describí en el capítulo 1 como una de tres estrategias de naturalización) y que este referente es el objeto de estudio.

En un momento posterior, se ofrece un candidato para tomar el lugar de dicho referente: el "conocimiento", tal como es usado como término teórico en etología cognitiva. La metodología

de la epistemología, de acuerdo con Kornblith, debe partir de que este es el fenómeno del que se habla al estudiar el "conocimiento." Dado que la etología cognitiva tiene un fuerte componente descriptivo —a diferencia de otras disciplinas que comparten el objeto de estudio, como la psicología comparada— Kornblith realiza una reconstrucción más precisa del conocimiento etológico en términos fiabilistas, ecológicos y evolutivos. Defiende la tesis de que dicho conocimiento es también el referente del conocimiento humano (argumentando contra objeciones internistas y contra otros tipos de externismo, como el fiabilismo tradicional de Goldman) y justifica su estatuto como clase natural apelando a una teoría de cúmulos estilo Boyd (como la expuesta en el capítulo 1). Finalmente, retoma un trabajo anterior sobre normatividad epistémica (1993) para defender por qué dicho conocimiento puede tener un papel normativo, además de ser valioso y por tanto deseable.

Quizá la única propuesta comparable en cuanto a estos criterios sea la de Ruth Millikan, quien se apoya en el marco desarrollado en sus trabajos anteriores para mostrar en un texto corto cómo se puede naturalizar el conocimiento (Cfr. "Naturalist reflections on knowledge", 1984). A pesar de que el aparato propio en el que se apoya Millikan es elegante, de gran poder explicativo y con presupuestos meta-filosóficos desarrollados de forma explícita, no hay una defensa propiamente epistemológica de la propuesta. El trabajo citado de Millikan contiene un mapa para aplicar las herramientas que ha desarrollado —particularmente la noción de función propia— para la epistemología, pero a diferencia de Kornblith no se enfrenta con posturas adversas en la literatura ni trata con detalle el problema de la normatividad epistémica. Lo anterior es natural, considerando que la epistemología no es su área de investigación principal. Por ello, al día de hoy, *Knowledge and its Place in Nature* es la exposición más completa de CN. Esta es la razón por la cual será la primer propuesta a considerar. Las propuestas estilo Millikan serán tratadas en el capítulo 3.

Acabo de resaltar algunas de las virtudes del planteamiento de Kornblith, pero desafortunadamente aunque el libro citado cubre la mayoría de los aspectos que se esperarían de una propuesta de naturalización, el desarrollo de los detalles es poco. Se trata de una elaboración sumamente esquemática, que ha sido aumentado en cuanto a sus alcances (pero no tanto en cuanto a detallar más la propuesta) en trabajos posteriores (e.g. Kornblith, 2007a, 2007b, 2010, 2011). Considero necesario realizar una reconstrucción de su propuesta con dos objetivos en mente: por una parte hacer explícitos los argumentos que utiliza para defender CN; por otra parte, hace falta descomponer los elementos de su propuesta para poder analizarlos con más detalle antes de ver cómo trabajan juntos. Esto llevará a su vez a profundizar más la interrogación, particularmente acerca de qué tipo de fiabilismo tiene Kornblith en mente. ¿Se trata de un fiabilismo en sentido estricto donde la fiabilidad está definida como una tasa de creencias verdaderas sobre creencias totales? Esto, como se verá en el capítulo, lleva a un problema importante, ya que plausiblemente una buena parte de lo que Kornblith considera como conocimiento natural no cumple con este requisito. Por otra parte, si la fiabilidad no se interpreta del modo tradicional, hace falta llenar vacíos teóricos que no están aclarados en la propuesta de Kornblith. Lo que sostengo en este capítulo, es que este problema nos debe llevar hacia otros desarrollos que no son simplemente versiones más detalladas de la propuesta de Kornblith, sino que llevan la tesis de la naturalidad del conocimiento en otras direcciones.

2.2 Conocimiento etológico

Para defender CN, Kornblith identifica un candidato en concreto: la noción de conocimiento utilizada por los etólogos cognitivos, que en adelante llamaré conocimiento etológico (CE). La noción de conocimiento es utilizada en trabajos de diversos investigadores como Ristau (1991), Heinrich (1999) y otros. A continuación describiré brevemente en qué consiste, enfatizando la

recepción de Kornblith sobre dichos trabajos. Antes, daré muy brevemente un contexto acerca de la etología cognitiva, como disciplina sucedánea de la etología.

La etología cognitiva es una recuperación de la etología clásica con la intención de integrarla al área interdisciplinaria que se empezó a denominar ciencia cognitiva, particularmente con la psicología comparada. La etología clásica de Konrad Lorenz, Nikolaas Tinbergen y Karl von Frisch surge en los años 1950s como una rama de la biología que compite con el conductismo dominante en psicología (incluyendo a la psicología comparada). Tomando como ejemplo el cognitivismo anecdótico de Darwin, la etología construye sus explicaciones partiendo de las atribuciones de estados intencionales a sus objetos de estudio (Allen y Bekoff, 1996). Lorenz va más allá de describir conductas animales en su medio natural a partir de un lenguaje mentalista e intenta postular explicaciones mecánicas⁵⁷ —con elementos como "energía motivacional"— para dar cuenta de los fenómenos sistematizados en sus trabajos previos. La deficiencia de las hipótesis tardías de Lorenz y la ausencia de una explicación más profunda sobre las regularidades que estudia son una de las razones de que la etología se enfoque entonces en proveer etogramas — descripciones de conducta animal en la naturaleza— más que estas explicaciones de carácter puramente especulativo. Hacia 1970 unos pocos etólogos como Lorenz y Hediger continúan hablando en términos de pensamientos y sentimientos (Griffin, 2001, p. 23). Otro factor que le quita relevancia a la etología clásica es que la misma psicología gradualmente se abre a un marco de trabajo mentalista.

Igualmente alrededor de esta década, el conductismo radical skinneriano ya no es la metodología dominante en psicología, sino un pluralismo metodológico (donde la conducta aún tiene un papel primordial) que permite postular representaciones "internas." Es en esta década cuando Donald

⁵⁷ Les llamo "mecánicas" en el sentido más general del término, no busco sugerir que sean materialistas, reduccionistas, etc.

Griffin (Cfr. *The Question of Animal Awareness*, 1976) busca darle otra dirección a la disciplina como etología cognitiva. La primera tesis metodológica que vale la pena mencionar, es que —al igual que los etólogos tradicionales— los etólogos cognitivos privilegian el estudio de los animales en su medio natural (permitiendo pequeñas intervenciones) y no en un laboratorio. No me adentraré en las razones para ello, ni sobre hasta qué grado sobre este punto las metodologías son rivales o complementarias; pero se puede encontrar un breve recuento en Burghardt (2013). En segundo lugar, la etología cognitiva en la tradición de Griffin es liberal cuando se trata de adscribir estados intencionales a los animales cognitivamente complejos.

En particular, Griffin —no sin controversia— es un defensor de la validez de apelar a la conciencia fenoménica e incluso en ciertos casos a la conciencia reflexiva [awareness], contra el consenso mayoritario (mas no absoluto) de los psicólogos cognitivos sobre estudiar funcionalmente la cognición siempre que sea posible. "[La] antipatía hacia considerar la conciencia amenaza con convertirse en una suerte de parálisis auto infligida, un lastre obsoleto para la investigación científica" (Griffin, 2001, p. 8). Contra esto hay una gran cantidad de programas de investigación en psicología cognitiva (incluyendo comparada) y etología, que utilizan las nociones de conciencia en un sentido únicamente discriminador (un ejemplo típico de uso discriminador es que el sujeto pueda decidir si una mancha es roja o azul, independientemente de su experiencia subjetiva). Dado que esto es un punto de disenso dentro de los mismos etólogos cognitivos, al igual que Kornblith, pasaré de largo el papel que pueda tener la conciencia en los estados intencionales (por ejemplo, si es una condición necesaria). En particular los estados intencionales que nos importan para intentar obtener una epistemología naturalizada a partir de la etología cognitiva son las actitudes proposicionales, y muy concretamente el conocimiento y la creencia.

Finalmente, un tercer aspecto metodológico que vale la pena resaltar —esta vez en contraste con la etología tradicional— es que la etología cognitiva está abierta a las conexiones interdisciplinarias. Tomemos un ejemplo, de acuerdo con Griffin puede ser útil apelar a trabajos en neurociencia

como la teoría del espacio global de Baars para postular mejores explicaciones acerca de lo que los animales realizan en un ambiente natural. En síntesis, etología cognitiva no es una disciplina tan aceptada ni con fundamentos tan sólidos como otras ciencias (e.g. la psicología cognitiva), aún para los estándares de las ciencias cognitivas, donde la relación entre distintas tradiciones de investigación —con distintos niveles de madurez— es compleja; en ocasiones existe colaboración, pero también hay enfoques complementarios y en competencia. No obstante, se trata de una tradición con una historia relativamente larga —tomando en cuenta sus orígenes en la etología clásica— y con una tendencia creciente a tener una relación complementaria con otro tipo de explicaciones. Al hacer notar esto, pretendo motivar que al menos es razonable adoptar como hipótesis de trabajo que la etología cognitiva pueda ser un comienzo para elucidar y posteriormente apoyar CN.

A continuación mostraré algunos pasajes que cita Kornblith para apoyar la idea de que en la etología cognitiva podemos encontrar un referente que realice el trabajo del conocimiento proposicional, el mismo del que se han ocupado los epistemólogos contemporáneos en las últimas décadas. En otras palabras, el "conocimiento" del que hablan los etólogos es el referente del término "conocimiento" que usan los epistemólogos. En adelante llamaré "conocimiento etológico" (CE) al primero. El siguiente fragmento citado por Kornblith (2002, p. 53) proveniente del trabajo de Ristau (1991) sobre cómo el frailecillo silbador (*charadrius melodus*) defiende su nido de depredadores ejemplifica el uso de este término:

En algunas ocasiones en que sea aproxima un intruso, este pájaro puede realizar una gradación de despliegues de ala rota, donde comienza abanicando la cola y gradualmente aumenta la torpeza de su caminar hasta que una o ambas alas están notablemente arqueadas, aleteando al tiempo que se arrastra [...] El pájaro presenta una actuación convincente de encontrarse herido y el observador se mueve cientos de metros para atraparlo sólo para verlo levantar el vuelo

ágilmente. para este momento el depredador ya se encuentra lejos del nido [...] [El frailecillo silbador] sabe cuándo un intruso es potencialmente peligroso.

La noción de conocimiento no es necesaria para explicar el comportamiento del frailecillo, ni el funcionamiento de su aparato cognitivo. La creencia basta para realizar dicha explicación, pues — al menos en una teoría representacional de la mente tradicional— las explicaciones se pueden realizar localmente (Pernu 2009; Fodor, 1987). El éxito de un frailecillo individual en una situación particular podría requerir ir más allá de su creencia para apelar a un componente fáctico: la creencia del frailecillo de que se acercaba un depredador era verdadera. Pero uno de los desiderata que indiqué en el primer capítulo es que conocimiento es más que creencia verdadera, una idea que suele acercarse al consenso en epistemología (con unas pocas excepciones de teóricos deflacionistas) y que Kornblith acepta. La creencia verdadera de que P es condición necesaria pero no suficiente para atribuir conocimiento de que P. Una vez que aceptamos que la creencia basta para explicar la conducta, y que la creencia verdadera basta para explicar el éxito, ¿qué explanandum queda que requiera postular el conocimiento? Kornblith considera que el explanandum no es el éxito de un individuo en una situación particular, sino el éxito adaptativo de los individuos pertenecientes a una especie, que requiere apelar a la historia evolutiva de sus capacidades cognitivas:

Para explicar por qué los frailecillos pueden proteger su nido tenemos que apelar a una capacidad de reconocer características del medio ambiente, y por tanto, a las creencias verdaderas que los frailecillos particulares adquirirán como producto de una capacidad estable para la producción de creencias verdaderas. Las creencias verdaderas resultantes no son accidentalmente verdaderas: son producidas por una capacidad afinada para un medio ambiente. [...] El concepto de conocimiento que interesa aquí requiere una creencia producida de manera fiable." (Kornblith 2002, p. 58).

Tal como lo interpreta Kornblith, el conocimiento es una clase natural remota, no próxima (Pernu, 2009). Si el conocimiento fuera una clase natural próxima, podríamos encontrarlo examinando el funcionamiento del aparato cognitivo de un conjunto de individuos del tiempo presente. En la ausencia de una historia evolutiva —por ejemplo, en el experimento mental clásico de que el mundo hubiera surgido hace cinco minutos, que se suele usar contra las teorías teleológicas de función— encontraríamos mecanismos "afinados" al medio ambiente, podríamos estudiar cómo trabajan, pero no deberíamos llamar "conocimiento" a sus productos. Si esto es el caso, no sólo es el caso que los mecanismos que producen el conocimiento etológico (CE) proceden de la historia evolutiva, sino algo más fuerte: la historia evolutiva es una parte *constitutiva* del CE. El CE no sólo es la capacidad de producir creencias fiables, afinadas al medio ambiente, sino que específicamente es una afinación/fiabilidad resultante de un proceso evolutivo.

2.3 Argumento metodológico

Identificar un candidato, en este caso el CE, para naturalizar el conocimiento es sólo un primer paso. Hace falta además, explicar en qué grado el CE naturaliza el conocimiento y argumentar por qué es el caso que esto es el caso, apelando a algún tipo de estrategia. Sobre el primer punto, la tesis de Kornblith es el caso extremo: el conocimiento *es* el CE. Es decir, en caso de ser exitosa, la naturalización es completa. Valdrá la pena problematizar esto más adelante. En cuanto al segundo punto, la estrategia que asume Kornblith para establecer su postura explícitamente es la que llamé de identificación de referente (cap. 1), aunque en un momento posterior defiende su teoría de objeciones mostrando que su naturalización del conocimiento cumple con las funciones epistémico-normativas (lo que se acercaría más a la estrategia de localización).

Debido a que Kornblith no se apega estrictamente a la estrategia de identificación de referente hace que sea necesario tender puentes entre el CE y la epistemología analítica tradicional. La

manera de hacer esto, es realizando una explicitación del CE en términos más familiares a la literatura filosófica. Esto lo trataré en el apartado acerca de la tesis epistemológica.

Volviendo al argumento metodológico, ¿en qué consiste entonces la estrategia? En términos muy informales la idea es que si una ciencia ya tiene una teoría acerca de un objeto o proceso del que hablamos en el lenguaje vulgar, debemos tener deferencia hacia dicha ciencia para decir en qué consiste el referente. El argumento recurre a un presupuesto semántico (ES) y tres presupuestos, que podemos llamar de naturalización (N):

ES. Externismo semántico para clases naturales.

N1. Compromiso ontológico con los referentes de los términos teóricos de una ciencia madura, en caso de que dichos términos cumplan ciertas condiciones.

N2. La etología cognitiva como ciencia madura.⁵⁸

N3. Presencia del CE en la etología cognitiva cumpliendo con los criterios de N1.

La tesis (a) es bien conocida, y corresponde a una familia de teorías del significado y la referencia propuesta por autores como Kripke (1980), Putnam (1975), y LaPorte (1997). Se trata de una postura contraria al descriptivismo, el internismo semántico más comúnmente sostenido. De acuerdo con el descriptivismo clásico sostenido por Frege, las ideas —públicamente accesibles— que tienen los hablantes acerca de las propiedades asociadas a un término determinan su significado, y éste a su vez determina la extensión. (Hay desarrollos donde el contenido del término puede ser psicológico personal en vez de abstracto; y versiones de la tesis donde se habla de un cúmulo de propiedades que sólo se debe cumplir en su mayoría). El descriptivismo aplicado a clases naturales tiene consecuencias inaceptables para caracterizar la actividad científica (Putnam

⁵⁸ O una ciencia cuyos métodos, categorías, etc, presumiblemente sobrevivirán en su versión madura o en una ciencia sucedánea madura.

1975, LaPorte, 2004). Tanto la idea de que los referentes de las teorías erróneas son vacíos como la discontinuidad de significado en la ciencia han sido consideradas como consecuencias inaceptables del descriptivismo (es decir, que el contenido semántico consiste en un conjunto de propiedades que los hablantes asocian al término y que dicho significado fija el referente⁵⁹), y adoptadas por tanto como razones fuertes para pensar que las clases naturales no se comportan como describe el internismo

Parecería entonces que una línea de objeción contra Kornblith sería cuestionar SE. Incluso si la semántica externista es la manera adecuada de pensar en el significado y la referencia de clases naturales, cabría dudar si los argumentos clásicos como los de Kripke (1980), formulados en términos esencialistas, aplican para las clases naturales boydeanas; que son las que plausiblemente pueden aplicarse a clases como el CE. No obstante, mostraré por qué no debemos centrar el interés en esta línea de crítica. Una mirada más cuidadosa revela que su argumentación no se apoya realmente en los aspectos propiamente semánticos, sino aquellos relacionados con la metodología: Kornblith no necesita adquirir compromisos semánticos tan fuertes para que marche su argumento. Por ejemplo, Tiel (1999) reconstruye el argumento de Kornblith apelando a que en la metodología científica los conceptos *antecedentes* —e.g. para agua, sustancia que está en los lagos, que se puede beber, etc. — se sustituyen a través de la práctica científica con nociones *consecuentes* (H₂O). Esto es una razón, *prima facie*, para pensar que el referente del concepto preteórico de conocimiento —si se trata del tipo de objetos que son de interés científico— será refinado por una teoría empírica. Considero que esta versión más débil de la premisa tiene sus problemas para salvar las atribuciones de conocimiento que realizamos de forma preteórica (puesto que si la semántica es descriptivista, nada garantiza que nuestra noción antecedente sea suficientemente adecuada para que determine la extensión que corresponderá al concepto consecuente). Como se

⁵⁹ El descriptivismo se ha formulado a nivel de conceptos psicológicos, pero también de "sentidos fregeanos", públicamente accesibles e independientes de la mente de cada hablante particular.

verá más adelante (cap. 4), las teorías con componentes evolutivos —entre ellas la de Kornblith— tienen problemas mucho mayores para salvar las atribuciones de conocimiento ordinarias. Por el momento dejaré esos problemas del lado, por lo cual concluyo que SE en su versión débil (metodológica) es una premisa altamente plausible, y que aún la versión fuerte es una buena premisa con más compromisos; pero también con un alto grado de plausibilidad.

El presupuesto N1, el compromiso ontológico con los referentes de los términos teóricos de una ciencia que cumpla con ciertas condiciones —las mencionaré en breve—, se deriva del realismo científico en su sentido más general. Este presupuesto tampoco es problemático, puesto que el debate sobre CN —como cualquier debate sobre si una clase es natural— suele presuponer una relación entre las clases naturales y los términos teóricos de las ciencias (notemos que cumplir este papel es uno de los desiderata de todas las definiciones de clase natural o de propiedades naturales que existen en la literatura).⁶⁰ Considero que las condiciones mínimas⁶¹ que deben cumplir los términos teóricos para tener compromisos ontológicos son, a grandes rasgos, ciertos criterios lingüísticos y de parsimonia. En primer lugar, resulta evidente que el uso de dichos términos debe ser literal, y no metafórico. No deben ser, de manera explícita o implícita, meros instrumentos heurísticos (esto excluye a las idealizaciones). En segundo lugar, en el caso de teorías formales debe haber claridad acerca de cuál es la interpretación pretendida (esto no es un problema para la etología cognitiva, que no es una disciplina matematizada). En tercer lugar, el término no debe tener un papel superfluo en la teoría (aquí podemos apelar a nociones de indispensabilidad), para lo cual es evidente que debe tener conexiones explicativas con otras partes de la teoría. Ir más allá de estas observaciones generales requeriría entrar en detalles sobre reconstrucción de teorías que no son relevantes para juzgar el presupuesto.

⁶⁰ En contraste, el naturalismo metodológico en epistemología no presupone un realismo científico. Al menos no se me ocurre un motivo a priori que impida al instrumentalista ofrecer razones para naturalizar la epistemología.

⁶¹ Otras condiciones que se pueden considerar son el poder explicativo y la convergencia. Agradezco a Dubian Cañas por hacerme notar esto.

La etología cognitiva como ciencia madura, N2, es el primer presupuesto realmente controvertido dentro de los parámetros del debate de naturalización del conocimiento. El punto fuerte de desacuerdo radica en el carácter intencional de las descripciones en etología cognitiva, como indica Bermudez (2006, 304): "Bien podría ser que la etología cognitiva haga un uso ineliminable de las nociones de creencia, deseo y conocimiento. ¿Pero esto muestra que el conocimiento es una clase natural o más bien que la etología cognitiva no es más ciencia que la psicología folk?" Por supuesto, la etología cognitiva no es la *única* disciplina que teoriza sobre la mente en términos intencionales, ni siquiera donde aparezca el conocimiento. Jennifer Nagel (2013) hace un recuento del uso del término "conocimiento" en diversas disciplinas, tales como la psicología comparada y la psicología del desarrollo, que parecen cumplir criterios como los que señalé en N1. Se podría defender que en disciplinas menos observacionales y con mayor tendencia a los controles experimentales, las descripciones antropomórficas tienen más oportunidades de ser refutadas —en contraste con lo que pasa en la etología cognitiva, llevada a cabo en ambientes naturales. Si esto no es suficiente para validar al CE —es decir, si no hubiera suficientes razones para considerar que CE sea el mismo referente del "conocimiento" de la psicología comparada, por ejemplo— en el peor de los casos, un seguidor del proyecto de Kornblith puede sustituir a CE por otra alternativa que cumpla con los desiderata del conocimiento que señalé en el primer capítulo.

Ahora bien, analicemos el siguiente escenario. Supongamos que existen diversas ciencias maduras que utilizan el "término" conocimiento C_1 , C_2 , C_3 . Todas las C's cumplen con los desiderata del conocimiento preteórico. ¿Se puede naturalizar en estas condiciones? Si no hay nada en común entre los C's más que cumplir con los desiderata, el conocimiento folk es una clase disyuntiva ($C_1 \vee C_2 \vee C_3 \dots$), y por tanto considero que se trataría o bien de un naturalismo eliminativista, pues no se conserva el conocimiento como clase, se sustituye por otras clases más precisas (como suele suceder con los términos folk), o en el mejor de los casos cuasi-eliminativista (si se mantiene el

trabajo teórico que realizaba el término preteórico). En este caso nuestras atribuciones de conocimiento vulgares podrían tener un uso instrumental al capturar un fenómeno de causalidad diversa (de manera parecida a lo que sostiene Dennett sobre el lenguaje intencional). Pero hay un mejor escenario. Quizá el conocimiento preteórico no es simplemente una clase disyuntiva —y por tanto artificial— sino que es un concepto que o bien captura ciertas generalidades que poseen todos los C's (de manera más fuerte que las que requiere un uso instrumental), o bien es un concepto cúmulo donde los C's comparten propiedades *más allá* de las desiderata iniciales (es decir, no sólo comparten propiedades epistémicas, sino también un suficiente número de propiedades psicológicas o biológicas, etc). Con lo anterior quiero ilustrar que el proyecto de elaborar un naturalismo no-eliminativista co-optando un término teórico como el CE (y otros similares), no es una vía cerrada, pero requiere de un trabajo muy detallado.

Como comentario al margen, aquí cabe replantear si el conocimiento realmente se puede reducir a una creencia con ciertas propiedades relacionales, o si estas C's nos obligan a adoptar un enfoque primitivista, donde el conocimiento sea irreducible a la creencia. Recordemos que Kornblith mantiene la ortodoxia de que la creencia tiene prioridad respecto el conocimiento, en contraste con el enfoque primitivista de Williamson (2000); aunque ambos tienen en común que defienden la tesis de que la epistemología se ocupa de objetos y no de conceptos; o dicho de otra manera, que el dominio de la epistemología no es metalingüístico (Brown, 2012), sino que refiere a algo concreto. Autores como J. Nagel (2013) y Engel (2006) hacen notar que una naturalización no-eliminativista del conocimiento buscando su identificación con algún C's, va de la mano con un espíritu primitivista (estamos dando por sentado que esas C's no se definen explícitamente apelando a la noción de creencia, en las respectivas disciplinas donde aparecen dichas C's). Las motivaciones de Kornblith posiblemente tengan que ver con evitar el problema de utilizar un estado mental fáctico para realizar explicaciones de la conducta de los individuos; a manera de ejemplo, Pernu (2008) hace notar los problemas semánticos sobre contenido mental sobre esta

línea de propuestas. Dejaré de lado este aspecto de la discusión, pero me pareció importante hacerlo notar.

Sobre N3, la presencia del CE en la etología cognitiva cumpliendo con los criterios de N1, ya indiqué anteriormente en qué consisten estos criterios. Es importante que CE sea utilizado cabalmente como un término teórico y que el uso del conocimiento sea literal en los escritos de los etólogos cognitivos que lo utilizan no sea un uso casual (como creencia verdaderamente). Esto no se ve tan claramente en el trabajo de Ristau anteriormente citado, pero en cambio aparece de manera explícita en el trabajo sobre delfines de Herman y Morrel-Samuels:

La competencias receptoras apoyan la adquisición de conocimiento, el bloque básico de un sistema inteligente. A su vez, el conocimiento y las habilidades de adquisición de conocimiento contribuyen de manera vital al éxito de un individuo en el mundo natural [...] Entre los requisitos básicos de un delfín adulto se encuentran las características geográficas y fisiográficas de su medio; las relaciones entre estas características físicas y rutas migratorias, la biota presente en el ambiente y su relevancia, [etc] [...] Esta lista es sin duda incompleta y parcialmente hipotética pero es ilustrativa de la amplitud y diversidad de la base de conocimiento necesaria para sostener la vida diaria del delfín [...] Se requiere un conocimiento extenso del mundo para funcionar de manera efectiva en el mundo y mucho del conocimiento requerido es ganado a través del ejercicio de habilidades receptoras" (Herman, Morrel-Samuels 1990, citado por Kornblith, 2002, p. 56)

Hay que tener en cuenta, además, los problemas a los que nos enfrentamos cuando las teorías que utilizamos como fuente de naturalización no son teorías formales (como en el caso de la física), o al menos con definiciones estandarizadas (como en mucho de la biología). En este caso, hace falta

un análisis cuidadoso al nivel metateórico del uso de dicho término⁶² y de los fenómenos que abarca, para comprobar que en realidad se trata de un término teórico robusto y conceptualmente estable. Es posible que en el camino encontremos obstáculos semánticos como vaguedad o sensibilidad al contexto, pero éstos no tienen por qué descartar de suyo la estabilidad conceptual o lo robusto del término. Sí conllevan, sin embargo, un trabajo de mostrar que se comportan bien desde el punto de vista teórico-metodológico a pesar de no ser términos unívocos.

Una vez teniendo claros los presupuestos (SE, N1, N2, N3), el argumento de Kornblith se puede reconstruir del siguiente modo. Señalaré entre paréntesis los presupuestos implícitos en cada premisa, haciendo explícita la diferencia de cuando el paso del argumento los tiene como presupuestos de cuando se deriva de ellos:

- 1) Los individuos pertenecientes a especies capaces de tener creencias tienen CE (premisa, presupone N1, N2, N3)
- 2) Por tanto, los humanos tienen CE (de 1)
- 3) Estrategia de fijación de referente: si el término "conocimiento" referido a los humanos fija como único referente a CE y CE es una clase natural, entonces el conocimiento es CE (premisa).
- 5) CE fija como único referente a CE (premisa, presupone SE, N1, N2, N3).
- 6) CE es una clase natural (premisa, se deriva de N1, N2, N3).
- 7) Por tanto el conocimiento humano es CE (de 4, 5, 6).

⁶² Jonatan García-Campos, en comunicación personal, me ha hecho notar que el proyecto de naturalización de la mente de Quine se apoya en la psicología conductista, que no es una teoría axiomatizada. Me parece que Quine en otros escritos provee análisis en una línea similar a la que describo (por ejemplo, en su análisis sobre el condicionamiento operante de Skinner, Cfr. por ejemplo Quine, 2013, p. 73). Por otra parte, a falta de una teoría axiomática un segundo nivel de precisión puede provenir de la operacionalización de los términos. Cabe notar que no todos los trabajos de etología cognitiva tienen un grado tan grande de operacionalización como el que existió en el conductismo o en ciertas áreas de la psicología cognitiva.

Kornblith no dice nada a favor de 1, quizá por considerarlo trivial. Quizá por razones evolutivas —o simplemente porque no tenemos noticia de casos semejantes— es difícil pensar en especies que posean un aparato cognitivo que o bien falle sistemáticamente en producir creencias verdaderas, o éstas sean siempre el resultado de una coincidencia fortuita. La interpretación que hace Kornblith del CE como un tipo de fiabilismo ecológico —trataré esto más adelante— es suficientemente amplia como para abarcar cualquier especie biológica que cumpla con las características señaladas. Lo que cabe cuestionar es qué tan problemática es dicha interpretación y qué tan adecuada es para caracterizar CE. Por ejemplo, ¿se podría seguir hablando de CE si las habilidades características de los seres humanos para adquirir creencias verdaderas se deben primariamente a la capacidad de aprender cosas (incluyendo métodos y técnicas) a partir de la reserva cultural como indica la hipótesis de inteligencia cultural (Herman y Austad, 2007)? Una razón para pensar que de hecho es así, es que esta capacidad de aprender requiere a su vez de competencias sociales innatas (ausentes en otros primates). Es decir, hay un conjunto de mecanismos evolutivos y de los cuales depende —en última instancia— la generación de creencias verdaderas.

La premisa 3 fue tratada en este capítulo y en el anterior. La premisa 6 fue tratada en este capítulo. Como señalé al evaluar los presupuestos N1, N2, N3, la derivación de estas premisas a partir de estos supuestos no es un asunto sencillo ni directo, pero mostré algunas formas como se pueden evitar las dificultades. Un aspecto importante es abrir al naturalista la posibilidad de una naturalización más amplia que incluya a CE, pero no exclusivamente; intentando que el resultado no sea una noción de conocimiento disyuntiva; lo cual conllevaría o bien que se trata de una clase artificial, o en el mejor de los casos, que el conocimiento es un fenómeno de causalidad dispersa. El asunto más problemático es la premisa 5, que el conocimiento humano fija como único referente a CE. En primer lugar, porque para ello es necesario realizar una explicitación de CE, intentando capturar de manera austera sus características. En la presentación de Kornblith este

paso no es explícito, pero su presentación de CE como teoría epistemológica consiste en una *interpretación* epistemológica que trataré en el próximo apartado. A partir de esta versión de lo que CE pretendidamente captura al nivel de generalidad que se busca en la epistemología, su exposición busca mostrar que su naturalización del conocimiento cumple con los requisitos de normatividad epistémica y que agota las adscripciones preteóricas de conocimiento. Esto implica descartar una tesis pluralista, donde CE sea una parte del conocimiento humano, pero exista otro conocimiento de distinto nivel o tipo (e.g. el conocimiento reflexivo) que no pueda reducirse a ello. De esto tratará el siguiente apartado.

2.4. Interpretación epistemológica del conocimiento etológico

La interpretación que realiza Kornblith de CE es lo que llamo un fiabilismo ecológico (FE). Por ello, la discusión se centrará sobre si FE agota el conocimiento humano; incluyendo la cuestión sobre si hay buenas razones para descartar el internismo epistémico frente a FE, un tipo de externismo: concretamente, de fiabilismo. Asimismo, será necesario ver si FE es capaz de responder a problemas epistemológicos a los que se enfrentan otros tipos de fiabilismo, como el fiabilismo de procesos de Goldman o las teorías de la virtud fiabilistas. La discusión puede llevarse a cabo por medio de contraejemplos y análisis de escenarios epistémicos prototípicos, pero es importante notar que esto no es necesariamente un regreso a la metodología tradicional de la epistemología, sino lo que expliqué en el capítulo 1 como una estrategia de acomodación. Como señalé entonces, Kornblith no es completamente ajeno a dicha estrategia, aunque oficialmente adopte una estrategia de fijación de referente.

Es necesario entonces hacer una aclaración sobre la metodología filosófica adoptada. No está de más resaltar una vez más que Kornblith no sólo está proponiendo un análisis del conocimiento,⁶³ al menos un análisis para ser evaluado por la comunidad filosófica de manera semántico-conceptual, aunque esto es lo que hace parcialmente al responder objeciones internistas (Cfr. 2002, cap. 2 y 3). La idea es que el estudio del conocimiento no debe ser acerca de nuestro concepto de conocimiento, sino que debe ser un estudio sobre el referente de dicho concepto. De acuerdo con esta visión, nuestro concepto preteórico o "antecedente" (Tiel, 1999) de conocimiento únicamente tiene la función de fijar un referente a partir de ciertos ejemplares, al modo como describen este proceso los trabajos de semánticas externistas (e.g. Kripke, 1980; Putnam, 1975). Así, Kornblith no defiende esta versión del conocimiento como una propuesta de análisis sujeto a contraejemplos, sino como una interpretación de la noción de conocimiento tal como aparece en la etología cognitiva. (Kornblith no hace estas distinciones, pero considero adecuado realizarlas para mantener abierta la posibilidad de que un examen más crítico pueda revelar que los usos de CE no siempre correspondan con FE).

Sugiero llamar fiabilismo ecológico-evolutivo (FE) a la caracterización hecha por Kornblith de lo que es el conocimiento etológico. La noción de conocimiento FE se puede descomponer de manera sencilla:

F1. S cree que P.

F2. P es el caso.

F3. La creencia de S de que P fue producida por un sistema seleccionado *para* generar una tasa adecuada de creencias verdaderas relativas al medio ambiente de S.

⁶³ Para una exposición detallada sobre las posturas metodológicas en epistemología de Kornblith ver 2002, cap. 1. Para un esbozo de una metafilosofía en la misma línea ver cap. 6.

F2 no tiene prácticamente controversia y simplemente representa el componente fáctico del conocimiento. Sobre F1, cabe notar que Kornblith comparte la visión extendida en epistemología de que el conocimiento es una propiedad que poseen cierto tipo de creencias (es decir, no es primitivo). Usualmente esto se expresa por medio de una prioridad conceptual, en un análisis semántico, pero generalmente (al menos en las teorías externistas) refleja una prioridad ontológica: la atribución de conocimiento "S sabe que P" refiere a un estado mental de S que posee ciertas propiedades además del componente fáctico (propiedades histórico causales, etc). Kornblith no sostiene que el conocimiento sea una clase natural "próxima" (Pernu, 2009), en el sentido de que el conocimiento no realiza un papel causal en las instancias que lo ejemplifican, pues esto lo lleva a cabo la creencia. Incluso cabría decir que sólo parte de que la creencia es la que realiza el trabajo causal, pues generalmente en las teorías sobre las actitudes proposicionales —con excepción de las teorías disposicionales— éstas poseen componentes que no realizan un trabajo causal (e.g. el componente histórico en las teorías teleológicas, la parte funcional en las teorías representacionales, etc.). Lo que realiza propiamente el trabajo causal son las estructuras mentales que sostienen este tipo de relaciones (e.g. los ejemplares de símbolos mentales en las teorías representacionistas). Por tanto, el conocimiento está reservado para animales cuyo aparato cognitivo sea lo suficientemente sofisticado para tener creencias.

Por lo anterior, es importante que CE sea utilizado cabalmente como un término teórico y que el uso del conocimiento sea literal en los escritos de los etólogos cognitivos, que no sea un uso casual (como creencia verdaderamente). Esto no se ve claramente en el trabajo de Ristau anteriormente citado, pero en cambio aparece de manera explícita en el trabajo sobre delfines de Herman y Morrel-Samuels:

La competencias receptivas apoyan la adquisición de conocimiento, el bloque básico de un sistema inteligente. A su vez, el conocimiento y las habilidades de adquisición de

conocimiento contribuyen de manera vital al éxito de un individuo en el mundo natural [...] Entre los requisitos básicos de un delfín adulto se encuentran las características geográficas y fisiográficas de su medio; las relaciones entre estas características físicas y rutas migratorias, la biota presente en el ambiente y su relevancia, [etc] [...] Esta lista es sin duda incompleta y parcialmente hipotética pero es ilustrativa de la amplitud y diversidad de la base de conocimiento necesaria para sostener la vida diaria del delfín [...] Se requiere un conocimiento extenso del mundo para funcionar de manera efectiva en el mundo y mucho del conocimiento requerido es ganado a través del ejercicio de habilidades receptivas (Herman, Morrel-Samuels 1990, citado por Kornblith, 2002, p. 56)

En este pasaje, CE parece ser algo más estrecho que FE, puesto que estos ejemplos de Herman y Morrel-Samuels sobre delfines, al igual que con los frailecillos de Ristau, son conocimientos de dominio específico. Esto es sumamente relevante, puesto que una de las tesis que defiende Kornblith es que el conocimiento humano y el conocimiento animal constituyen una sola clase. Pero resulta plausible —aunque no universalmente aceptado— que al menos una parte importante del aparato cognitivo humano responsable de formar creencias no es de dominio específico. Por el momento no entraré a mayor detalle sobre este punto, en parte porque esto requeriría un estudio detallado de la literatura de etología cognitiva, pero es importante tener presente que si la caracterización de Kornblith no coincide de manera suficiente con el uso del término "conocimiento" en etología cognitiva esto abre una serie de problemas.

El conocimiento etológico en el fragmento de Herman y Morrel-Samuels y otros similares, realiza un trabajo explicativo que no podría hacer la noción de creencia. La creencia explica el comportamiento de individuos, mientras que CE de acuerdo con Kornblith explica dos cosas: el éxito sistemático de los individuos de la especie y la existencia del aparato cognitivo de dicha especie en su forma presente. Agregando algunos supuestos relacionados con argumentos de utilidad teórica y sobre clases naturales, la función explicativa de CE es evidencia *prima facie* de que CE existe y de que CE es una clase natural. Además de que un análisis de CE/FE puede mostrar

que se trata de una clase natural, que para Kornblith sería la teoría de Boyd de cúmulos de propiedades causalmente homeostáticas mencionada en el capítulo anterior.

Cabe preguntar ahora si las atribuciones de conocimiento en humanos tienen como referente —y más aún, como único referente— al CE/FE. Las posibilidades a considerar son las siguientes:

a) Identificación reductiva: Las adscripciones de conocimiento refieren simplemente a creencias que son CE.

b) Identificación reductiva parcial: Las adscripciones de conocimiento refieren a creencias que son CE y que además tienen otras propiedades (e.g. reflexión, coherencia, articulación social, etc). En este caso, lo que hace a una creencia verdadera conocimiento son las propiedades Φ y Ψ , etc, pero sólo Φ se puede identificar con una propiedad natural. Es decir, CE es una condición necesaria pero no suficiente para adscribir conocimiento.

c) No-identificación: Las adscripciones de conocimiento refieren a creencias que no son CE (excepto de manera accidental o coincidente), y que son conocimiento debido a otras propiedades (e.g. reflexión, coherencia, etc; pero también fiabilidad —en sentido tradicional—, origen causal, etc.). Es decir, CE no es conocimiento (en el sentido epistemológico relevante) ni una condición necesaria para el conocimiento.

d) Referentes múltiples: Las adscripciones de conocimiento pueden tener como referencia a creencias que tienen las características descritas en a, b o c (por lo menos dos).

Resulta claro que la posición que defiende Kornblith es la primera. Contra internistas como Bonjour (2006), sostiene que la reflexión sobre la evidencia no tiene ningún valor por sí mismo. Kornblith refuta argumentos internistas en dos ámbitos principalmente: la reflexión, como un requisito de segundo orden, y la necesidad de una sociedad actual o posible contra la cual contrastar y defender las creencias a la manera que defienden Williams, Brandom y Kusch. No vale la pena

reseñar aquí ambas discusiones, puesto que realmente son las discusiones tradicionales de internismo vs externismo / individualismo vs no-individualismo⁶⁴, sin que el naturalismo no-eliminativista de Kornblith tenga algún papel en particular. La conclusión de Kornblith, es que el proceso de reflexión por sí mismo no provee algún valor especial a las creencias verdaderas que las convierta en conocimiento, en otras palabras "no hay magia de segundo orden" (Kornblith, 2010). Vale la pena recordar que en este contexto por reflexión se entiende algo muy débil, simplemente que el agente confronte una creencia con la evidencia disponible que posee y/o con el resto de su conjunto doxástico. No se trata de un proceso de reflexión a priori, que implique certeza y/o necesidad epistémica. Teniendo esto presente, Kornblith cita ejemplos de experimentos psicológicos (Nisbett y Ross, 1980) donde se muestra que la reflexión de los sujetos del estudio sólo tiene como resultado incrementar el nivel de confianza en sus propios juicios erróneos sin llevar a corregirlos, poniendo en duda que la consideración reflexiva tenga algún valor en sí misma; de hecho en el estudio citado es perjudicial.⁶⁵

Ahora bien, Kornblith no sostiene que la reflexión nunca pueda proveer valor, sino que en ciertas ocasiones lo hace pero porque aumenta la fiabilidad. Ante este tipo de evidencias, el internista se encuentra presionado a restringir el tipo de consideración reflexiva que debe ser adecuada, como postular que en el largo plazo y con una mayor base de creencias los sujetos podrían percatarse de su error. Esto genera dos problemas para un internista. En primer lugar —por simple cálculo de probabilidades— conforme aumentan las exigencias, la cantidad de creencias que merecen el título de conocimiento tiende a ser menor.

⁶⁴ Distintos autores no-individualistas usan términos diferentes para nombrar su postura. Por ejemplo, Kusch denomina a su teoría comunitarista, mientras que Brandom se autodenomina inferencialista social, etc.

⁶⁵ Bonjour (2006, p. 323) rechaza la validez de estos contraejemplos particulares, pues 323 "no involucran el tipo de inferencia explícita y consciente que según el internista debe estar disponible para la introspección." No obstante, me parece que esto deja la puerta abierta a nuevos experimentos sobre la fiabilidad de la introspección, buscando maneras de garantizar que el sujeto reflexione cumpliendo las condiciones impuestas por los internistas.

El internista siempre puede agregar nuevos requisitos, pero conforme lo hace su objeto de estudio cada vez se aleja más del conocimiento proposicional ordinario.⁶⁶ El segundo problema para el internista radica en que un fiabilista podría responder que la razón por la cual Bonjour pide —por ejemplo— que los resultados de la reflexión se vean a largo plazo, es precisamente porque en el largo plazo la reflexión presumiblemente sería capaz de proveer la fiabilidad que no provee en las condiciones del experimento. En resumen, resulta sospechosa la insistencia del internista de que la consideración reflexiva daría buenos resultados en otras condiciones, no porque esto resulte implausible, sino porque el énfasis en los buenos resultados da razón para pensar que este es el criterio en última instancia. Es imposible reconstruir aquí de manera justa el debate entre internismo y externismo. Una valoración global requeriría entrar de manera detallada en problemas como el regreso epistémico.

¿Qué hay del reduccionismo parcial? En diversos trabajos (e.g. 2009, 2010, 2011), Sosa distingue entre dos niveles: el conocimiento animal y el conocimiento reflexivo. El primero está planteado a partir de la epistemología de la virtud, que es un externismo centrado en el agente. Incluso — aunque no comparto del todo esta opinión— se le suele considerar un tipo de fiabilismo ("El conocimiento animal [de Sosa] es a grandes rasgos, creencia verdadera producida de manera fiable.", Kornblith 2004, p. 126). El análisis de Sosa, basado en la tesis del conocimiento como logro [knowledge as achievement] es el siguiente: una creencia es conocimiento animal si y sólo si es verdadera, fue adquirida de manera virtuosa o adecuada y es verdadera porque fue adquirida de manera adecuada. No obstante, es posible tomar la idea de que existen estos dos niveles de

⁶⁶ Por ejemplo, un requisito de responsabilidad epistémica difícilmente sería compatible con la atribución de conocimiento a infantes, a casos donde los sujetos no recuerdan la evidencia en que se apoyaron para generar una creencia o a adquirir conocimiento a partir de testimonios que se aceptan *prima facie* como verdaderos.

conocimiento sin aceptar este análisis en particular. Cabe señalar que Sosa no suele ser explícito con respecto a la parte metafísica, sobre si el conocimiento animal es una clase natural.

El primer nivel de conocimiento es el que, como su nombre lo indica, compartimos con los animales no-humanos que son sujetos de atribuciones epistémicas. El conocimiento reflexivo toma como punto de partida las creencias que constituyen al conocimiento animal, pero agrega un examen del propio agente acerca de la evidencia que existe para dicha creencia así como la coherencia con otras creencias dentro del conjunto doxástico del agente. Los ejemplos que provee Sosa (2009) son muy convincentes. Uno de ellos consiste en los siguientes casos:

- a) Una persona observa una superficie de color rojo en condiciones de luz normales con su aparato visual y cognitivo en buen estado y adquiere la creencia verdadera de que la superficie es roja sin mayor reflexión.
- b) Igual que el escenario anterior, excepto que la persona además tiene creencias adicionales acerca de las condiciones en que fue producida dicha experiencia.

Es claro que en el primer caso podemos atribuir conocimiento animal y en el segundo caso, tanto conocimiento animal como reflexivo acerca de lo mismo (el color de la mesa). El externista negaría que tener conocimiento reflexivo sobre el color de la mesa esté aportando algo con respecto a tener conocimiento animal. Sosa retoma estos escenarios, pero en un contexto donde hay un bromista que suele iluminar las mesas y cambiarlas de color; sólo que dicho bromista decide no hacerlo. Esto conlleva que la creencia verdadera del sujeto se debe en parte a la suerte epistémica: el bromista fácilmente podría haber decidido usar sus reflectores y hacer que el sujeto generara una creencia falsa. En estas condiciones, puesto que el aparato cognitivo y visual del sujeto funciona de manera virtuosa, es posible tener conocimiento animal sin saber de la existencia del bromista, pero para tener conocimiento reflexivo es necesario descartar que el bromista esté

interviniendo. En este caso no se trata simplemente de que la reflexión aumente la fiabilidad de la creencia (aunque de hecho lo hace), sino que al descartar la suerte epistémica provee un valor adicional.

El carácter burlesco de este experimento mental no debe oscurecer el hecho de que este tipo de escenarios no son extremos. No es difícil encontrar ejemplos más realistas. Por ejemplo, los pasajeros (y la tripulación) de un avión con determinados movimientos sufren de ilusiones vestibulares, por ejemplo, sentir que el avión mantiene su trayectoria cuando en realidad está descendiendo. Durante un vuelo, las ilusiones del sistema vestibular no son constantes, de manera que la percepción del movimiento del avión que tiene uno de sus ocupantes la mayoría del tiempo es adecuada (y por tanto sus creencias generadas a este respecto). Resulta difícil negar que un pasajero que adquiere creencias verdaderas sobre la orientación del avión debido al buen funcionamiento de su sistema vestibular tiene —al menos en alguna medida— conocimiento. Al mismo tiempo, resulta igualmente difícil negar que un pasajero capaz de descartar que el avión esté realizando una maniobra que pudiera fomentar una ilusión vestibular se encuentra en una situación epistémica considerablemente mejor.

Un fiabilista de proceso que no estuviera comprometido con la naturalidad del conocimiento, podría simplemente responder que sólo hay un tipo de conocimiento y que el pasajero del avión no sabe que P (donde P es una proposición sobre la dirección del avión) a pesar de que cree que P y P es verdadera. Esta salida no está disponible para un defensor de FE, dado que el sistema vestibular está funcionando de manera adecuada en ese momento. Lo más fácil sería admitir que la situación epistémica del pasajero reflexivo es mejor, pero que esto no se debe a la existencia de dos niveles de conocimiento.

Una posibilidad es que el pasajero reflexivo tenga un conocimiento animal de la posición del avión y que además su creencia tenga mayor justificación que la del pasajero irreflexivo (sin que esto requiera postular un nuevo tipo de conocimiento). Otra alternativa sería explicar la mejoría en la situación epistémica señalando que el segundo sujeto conoce dos proposiciones distintas, una de segundo orden, Q, sobre su conocimiento de primer orden. Aquí la explicación concreta dependerá de cómo se llenen los detalles del escenario. Si es un pasajero al que se le informó la orientación del avión, el conocimiento de Q se puede explicar por la capacidad de aceptar un testimonio confiable. Si se trata de un piloto que puede ver su instrumental, el conocimiento de que Q puede fundamentarse en su capacidad para leer los números y manecillas y en su capacidad para realizar razonamientos inductivos; o dado que el piloto recibió una instrucción, también apelando al testimonio. Con este tipo de estrategias, se puede ver que Kornblith no necesita insistir en desestimar el papel de las creencias de segundo orden como fuente de valor epistémico, sino simplemente aceptar que existe este valor y proveer explicaciones alternativas; por supuesto sin aceptar la tesis internista de que los requisitos de segundo orden sean una condición necesaria para el conocimiento.

Como mencioné anteriormente, la caracterización del conocimiento que realiza Kornblith (FE) se distingue por su tercer elemento: F3 La creencia de S de que P fue producida por un sistema seleccionado *para* generar una tasa adecuada de creencias verdaderas relativas al medio ambiente de S. El objetivo de esta sección es ver con mayor detalle el contenido de F3. Por razones de exposición centraré la discusión en el aparato cognitivo humano, pero lo dicho en esta sección es igualmente aplicable a los animales no-humanos capaces de tener creencias.

Lo primero que hay que notar, es que F3 tiene dos elementos. Por una parte, hay un planteamiento fiabilista (la tasa de creencias verdaderas) y por otra parte, un elemento teleológico (la referencia a la selección natural). El fiabilismo sostenido por Kornblith es muy parecido al fiabilismo sobre el

conocimiento, como el de Goldman (1988, basado en su trabajo previo sobre la justificación de 1979); lo cual notan explícitamente ambos autores (Kornblith 2002, p. 63-69; Goldman, 2005).

Sin entrar en detalles, la idea fundamental de Goldman es mantener el análisis tradicional del conocimiento (creencia verdadera justificada), con una noción de fiabilidad que hace el papel de justificación. La fiabilidad se define con base a una tasa de creencias verdaderas sobre falsas que daría un proceso (incluyendo, pero no exclusivamente, procesos cognitivos) en un conjunto de escenarios o mundos posibles relevantes. (Dado que los mundos posibles en la semántica estándar tienen una cardinalidad incontable esto genera algunos problemas técnicos que se pueden resolver partiéndolos en clases de equivalencia). Sin tomar en cuenta problemas que se han planteado posteriormente al fiabilismo como el problema de la clase de referencia, hay dos cuestiones importantes por explicitar: qué grado de fiabilidad se requiere para atribuir conocimiento y cómo se determina cuáles son los mundos relevantes (por ejemplo, descartando escenarios escépticos). Ambas cosas están dadas por la semántica del término conocimiento.

Dado que Goldman sostiene que los términos epistémicos son clases artificiales —primero de manera implícita en 1986, posteriormente de manera explícita en 2007— no hace falta apelar a algún elemento no arbitrario que resuelva estas dos cuestiones. Simplemente los agentes epistémicos compartimos un concepto de conocimiento que considera ciertos escenarios como relevantes y que exige cierto grado de fiabilidad.

En ciertas ocasiones Kornblith (2002, p. 63) plantea el FE como una versión del fiabilismo que por una parte sustituye a los procesos por capacidades cognitivas del agente (de manera parecida al conocimiento animal de Sosa, Cfr. 2009) y por otra parte, determina de un modo objetivo los dos puntos anteriores. Los escenarios relevantes están dados por el medio ambiente para el cual fueron seleccionadas dichas capacidades cognitivas, mientras que el grado de fiabilidad es el

necesario para que dichas capacidades cognitivas fueran seleccionadas. Kornblith hace notar que lo anterior resuelve problemas como la inestabilidad en las atribuciones de conocimiento en casos como las fachadas de granero de Brandom (1998). Aunque para Brandom dicha inestabilidad no es problemática, ésta es una consecuencia que para Goldman (1993, p. 163) es inaceptable.

También resulta plausible —aunque no lo examinaré aquí— que FE tenga recursos para responder mejor a los problemas que han sido planteados al fiabilismo de proceso. El caso del problema de la generalidad (Cfr. Conee y Feldman, 1998) es notorio, dado que ha sido particularmente persistente para las distintas versiones del fiabilismo. En resumidas cuentas, este problema indica que un ejemplar de proceso puede caer bajo distintos tipos con distintos grados de fiabilidad, el fiabilismo tal como fue planteado por Goldman (1979) no da elementos para señalar cuál de estos tipos es el correcto. Una teoría fiabilista de proceso no puede ser completa sin resolver esta cuestión.

La propuesta de Kornblith presumiblemente resuelve esta determinación, pero al mismo tiempo se aleja del espíritu de simplicidad del fiabilismo de proceso, al introducir elementos teleológicos que conllevan un elemento histórico evolutivo. ¿Sería posible constreñir el fiabilismo de proceso con un espíritu naturalista sin incurrir en costos teóricos tan altos —con los problemas que conllevan— como el FE?

Como ejemplo, está la solución propuesta por Alston (1995, p.5), quien señala que "hay consideraciones fundamentales que marcan, para cada ejemplar de proceso, un tipo que es su "clase natural." ¿Cuál es el problema con esta propuesta? Aun haciendo la restricción a clases naturales, los ejemplares de proceso pueden caer bajo distintos tipos. En palabras de Conee y Feldman (1998, p. 145):

los ejemplares de proceso pueden pertenecer a clases naturales. No obstante, no hay una razón para pensar que pertenezcan a una sola clase natural [...] Cada ejemplar de proceso de formación de creencias es categorizado de distintas maneras por diversas ciencias [...] Los candidatos razonables para clases naturales para un proceso típico de formación de creencias visuales incluye procesos electroquímicos, orgánicos, perceptuales, visuales y de reconocimiento de rostros [...] Estas clases naturales difieren enormemente en su grado de fiabilidad, de manera que [...] [no queda resuelto] el problema de la generalidad.

Restringir la partición de procesos a clases naturales no es suficiente para determinar cuál es el tipo de proceso relevante en cada ejemplar. La propuesta de Kornblith —aunque nunca lo plantea de esta manera— podría verse como un modo natural de extender la respuesta de Alston, introduciendo más elementos que permitan identificar un ejemplar de proceso con una sola clase natural relevante en cada caso. La manera de hacer esto es con una referencia a la capacidad o competencia que produjo la creencia y que fue seleccionada *para* ello (lo cual a su vez explica la existencia de dicha competencia). El FE queda libre del problema de la generalidad gracias a los elementos teleológicos que introduce.

Sin embargo, una vez introducido el elemento teleológico, resulta hasta cierto punto contraproducente limitar las explicaciones teleológicas con un requisito de fiabilidad. Como señala Jessica Brown (2012, p. 43) contrastando el proyecto de Kornblith con las noción de función propia de Millikan:

Un determinado sistema perceptual de un roedor podría tener la función propia de detectar halcones, aún si fuera poco fiable y diera muchos falsos positivos, siempre que su presencia continua en los miembros de dicha especie de roedores fuera parcialmente explicada por el hecho de que contribuye a que los miembros de la especie eviten a los halcones.

De acuerdo con Kornblith, el CE/FE es necesario desde el punto de vista explicativo para dar cuenta de la existencia de los mecanismos que producen las creencias que ameritan ser llamadas

"conocimiento", pero como muestra Brown, la fiabilidad al momento de llevar a cabo dicha explicación. Esto arroja más dudas acerca de que FE sea una buena interpretación de CE. Al menos en los ejemplos citados por Kornblith no hay nada que haga pensar que los mecanismos que generan una gran cantidad de falsos positivos no caerían bajo la clasificación de "conocimiento" de los etólogos cognitivos. Suponiendo que Kornblith admitiera esto, aun así podría decidir restringir el conjunto de creencias que son conocimiento a aquellas que cumplan con el requisito de fiabilidad.

El problema sería de qué manera se puede justificar este movimiento, pues si acepta que FE sólo es un subconjunto de CE, la existencia de CE como término teórico ya no es suficiente para sostener que FE es la referencia del término "conocimiento." Una vez más, cabe insistir que la estrategia metodológica de Kornblith no consiste en proponer un análisis para ser evaluado de manera semántico-conceptual. También esto genera dudas acerca de que la caracterización de Kornblith sea una clase natural en el sentido de Boyd. Si la fiabilidad es innecesaria en el sentido anteriormente descrito, no parece que FE sea un cúmulo de propiedades causalmente homeostáticas. El problema todavía es más grande si, como señala Millikan, además de que dichos mecanismos no sólo pueden no ser fiables, sino que regularmente no lo son:

La caracterización de Millikan de función propia plantea la posibilidad de que al ver los mecanismos psicológicos de formación de creencias a través del lente de la evolución al modo como busca Kornblith, uno pueda terminar con una caracterización de funciones propias que con frecuencia no se ejecutan. De hecho Millikan (1984, pp. 59-70) explícitamente afirma que es probable [likely] que las funciones propias de los mecanismos psicológicos y los estados que producen queden frecuentemente incumplidas, dado que su cumplimiento depende del mundo que está más allá del organismo [...] La discusión de Millikan sugiere que un enfoque darwinista a la taxonomía de los mecanismos que producen creencias podría no enfocarse en la noción de creencia fiablemente producida, sino en la categoría de un estado que se dirige a la verdad. (Brown, 2012, p. 43)

La tesis de que la selección natural produce mecanismos fiables en relación a las creencias verdaderas que producen dichos mecanismos es un lugar común en la filosofía (Stein., 1996); como se puede ver en el caso de Dennett (1978), Fodor, (1981), Goldman (1986), Quine, (1969b). La idea es que para las especies que poseen capacidades cognitivas, dichas capacidades son una adaptación, por lo que son representacionalmente adecuadas. ¿Cómo debemos interpretar este requisito? Al menos existen dos maneras:

- a) La competencia fue seleccionada por/para ser fiable.
- b) La competencia fue seleccionada y es fiable.

Como se verá en el siguiente capítulo, el requerimiento de que una competencia haya sido seleccionada por/para (ambas alternativas conllevan distintas cosas) ser fiable presenta problemas importantes. Pero incluso el requerimiento mismo de fiabilidad resulta altamente problemático. El problema de los falsos positivos no parece ser meramente una excepción cuando se estudian fenómenos naturales como el CE, sino algo prevalente. Un naturalista estilo Kornblith puede insistir en sostener b, señalando que su objetivo es obtener los beneficios del fiabilismo evitando los problemas de subdeterminación del fiabilismo de procesos —el problema de la generalidad— o del fiabilismo de virtud —la ausencia de una semántica clara acerca de qué tipo de procesos son virtuosos, y las trabas de naturalizar una versión del conocimiento donde la normatividad juega un papel constitutivo. Pero la partición de Kornblith es un subconjunto de los mecanismos estudiados por el CE —y plausiblemente, por otras nociones de conocimiento $C_1, C_2, C_3 \dots$ — y no hay una razón por la cual hacer esa partición. Un fiabilista que no es naturalista ontológico —al menos no completamente— como Goldman, puede sostener que la fiabilidad es escogida debido a que la metodología epistemológica busca refinar la noción de conocimiento folk, es decir, rechaza CN. Esto es compatible con que las ciencias cognitivas puedan ayudar a especificar los escenarios relevantes donde importa la fiabilidad (con el presupuesto de que la noción de fiabilidad captura

escenarios "normales" y que las ciencias cognitivas se ocupan de estos mismos escenarios normales). Aquí es útil notar que Goldman (2005) manifieste simpatía metodológica por utilizar a disciplinas como la etología cognitiva para establecer los escenarios relevantes, pero rechaza explícitamente el compromiso con CN.

2.5. Conclusión del capítulo

¿Qué debemos entender por un término teórico? Me parece que sería erróneo tratar de escoger un criterio de teoreticidad (por ejemplo, comparando con un lenguaje observacional) para posteriormente confrontar la literatura citada por Kornblith. Lo que le da peso a la segunda premisa no es que t sea teórico (lo que sea que eso signifique), sino que t —como propiedad de orden superior adscrita a determinadas creencias— se utilice de manera explicativa y proyectiva. En palabras de Antony (2003) —en el contexto de otra discusión— "hay un repertorio causal asociado con dicha propiedad."

Es necesario examinar entonces tanto el uso explicativo como el uso proyectivo de CE. Me enfocaré aquí únicamente en lo primero. Kornblith (2002, p. 57) señala lo siguiente:

Cuando pasamos a una explicación de las capacidades cognitivas de una especie, la empresa teórica en la que estamos involucrados requiere más que mera creencia. Ya no estamos interesados en explicar por qué un frailecillo particular se movió de su nido en un modo que atrajo la atención del depredador; sino en una explicación de por qué los miembros de su especie tienen una capacidad cognitiva que les permite interactuar [negotiate] de manera exitosa con el medio ambiente. El enfoque en la adaptación de estas capacidades cognitivas para un medio ambiente que nos obliga a explicar la posibilidad de comportamiento exitoso es lo que requiere la noción de conocimiento en vez de mera creencia. El conocimiento explica la posibilidad del comportamiento exitoso en un medio ambiente, que a su vez explica la aptitud.

¿Qué es lo que busca decir Kornblith aquí? ¿Que el explanandum de la etología cognitiva sería primariamente la posibilidad de conducta exitosa e indirectamente la aptitud? Lo segundo resulta implausible. La aptitud genética no aparece como explanandum en los escritos de etología cognitiva. Quizá una lectura más caritativa señalaría que la existencia de un rasgo cognitivo —un explanandum de la historia evolutiva— se explica por medio de la aptitud, y que ésta a su vez se explica por medio del papel adaptativo que tienen dichas capacidades. En este caso, ambos explananda (la posibilidad de conducta exitosa y la aptitud) tendrían un carácter independiente, y de este modo cada una daría apoyo por separado a la función explicativa y por tanto a la utilidad teórica de CE.

Una vez que se identifica a un tipo de conducta, las nociones intencionales explican y unifican dicha conducta. Bermudez (2006) hace notar que la etología cognitiva es idiosincrática entre las ciencias cognitivas por proceder de esta manera, cuando el enfoque de las ciencias cognitivas exitosas no es explicar conductas sino la función de mecanismos (Cummins, 2000). Contra esto, se pueden ofrecer numerosos ejemplos en otras disciplinas (ver J. Nagel, 2013 para una lista parcial del uso del término "conocimiento" en psicología cognitiva y psicología del desarrollo). Suponiendo que nociones intencionales como la creencia y el deseo no sean problemáticas, aún cabe preguntarse por qué la conducta *exitosa* (y no sólo la conducta) es un explanandum de la etología. Por ejemplo en el caso de los delfines de Herman y Morrel-Samuels (1990), de los cuales se postula que tienen las competencias receptivas necesarias para realizar un cierto tipo de conducta (navegar adecuadamente en su medio ambiente).

¿Cuál sería la diferencia si se tratara de explicar una conducta maladaptativa? Frecuentemente las conductas maladaptativas se enfocan en individuos —no en especies— apelando a fallas mecánicas o problemas ontogenéticos. Pero dado que la evolución no es óptima, también existen conductas maladaptativas a nivel de la especie (para ejemplos ver Dawkins, 1976). Plausiblemente algunas de

estas conductas pueden explicarse con un enfoque cognitivo, es decir, a partir de representaciones inadecuadas. Pero no es necesario apelar a una historia evolutiva para dar una explicación próxima de la conducta. Al tratarse de un fenómeno biológico, si pedimos explicaciones de manera iterativa eventualmente llegaremos a la historia evolutiva de dicha especie, pero esto no implica que dicha iteración sea necesaria para dar con un explanans satisfactorio.

El nivel de profundidad de las explicaciones etológicas parece satisfacerse postulando representaciones y mecanismos representacionales, sin que la historia evolutiva sea estrictamente necesaria; es útil recordar que la etología cognitiva a partir de su origen en los años 70 por investigadores como Donald Griffin se enfoca en el modo como los animales "perciben, interpretan y experimentan el mundo" (Rivas y Bughardt, 2002). Las explicaciones de conducta adaptativa en principio no tendrían que ser distintas en este nivel. Por estas razones, cabría pensar que el uso del término "conocimiento" en etología cognitiva no tiene una función explicativa indispensable. No hay ningún problema en decir que los frailecillos *saben* cuándo se acerca un depredador, pero este uso del término "saber" es intercambiable por otro (e.g. los frailecillos *detectan* cuando viene un depredador). Si hay un explanandum de representaciones y conducta exitosas, habría que buscarlo a nivel de la biología evolutiva.

El problema radica en que en la biología evolutiva no hay un explanandum que abarque todas las representaciones exitosas de los animales con capacidades cognitivas o siquiera de una sola especie. La existencia de cada mecanismo —pueda o no individuarse como rasgo— requiere de una explicación histórica respectiva; como cuando un geólogo explica por qué una placa terrestre se encuentra en determinada posición. Ahora bien, a pesar de que estas capacidades adaptativas no constituyen un sólo explanandum, lo cual es otro impedimento más al argumento metodológico de Kornblith

He mostrado en este capítulo que tanto el fiabilismo como el compromiso con defender su postura con un argumento metodológico dificultan la postura de Kornblith. Una teoría que deje de lado la fiabilidad y considere al conocimiento en términos puramente teleológicos tiene mejores prospectos para resolver los problemas mencionados en este capítulo. En el siguiente capítulo desarrollaré esta línea de naturalización. No obstante, mostraré que la versión puramente teleológica de naturalización presenta problemas insolubles. Aunque no continuaré el diálogo con Kornblith, el lector podrá comprobar en el capítulo 3 que los problemas derivados de apelar a la historia evolutiva para naturalizar el conocimiento se aplican igualmente a la propuesta criticada en este capítulo.

CAPÍTULO 3: EL CONOCIMIENTO COMO FUNCIÓN PROPIA

3.1. Introducción

En el capítulo anterior se estableció la idea de que el fiabilismo ecológico (FE) de Kornblith tiene problemas irresolubles. En resumen, los mecanismos cognitivos no surgen debido a que produzcan una tasa mayor de creencias verdaderas, sino a que producen aptitud. Dicha aptitud está en función de otros elementos como el uso de recursos (cognitivos y de otro tipo, como energéticos) y distintas funciones de utilidad de creencias que no tienen que ver exclusivamente con el número de creencias falsas sobre verdaderas, como se puede ver en el ejemplo de los falsos positivos. Como ejemplo, un mecanismo cognitivo para el comportamiento antipredatorio puede generar mayor aptitud produciendo una gran cantidad de falsos positivos, es decir, si es extremadamente sensible.

Una segunda área problemática de la propuesta de Kornblith es la identificación del conocimiento con el término técnico utilizado en la literatura de etología cognitiva (conocimiento etológico o CE). La preocupación es que dicha identificación pueda dar como resultado una visión demasiado estrecha del conocimiento como para que pueda ser una teoría fructífera del conocimiento humano.

La solución propuesta fue abandonar la fiabilidad como criterio y hacer una naturalización del conocimiento centrándose exclusivamente en la idea de función, abandonando la idea de fiabilidad como un elemento constitutivo en la definición del conocimiento. Lo anterior, por supuesto, no implica que la aparición de mecanismos cognitivos fiables no sea una consecuencia frecuente de la selección natural.

La tesis del conocimiento como función propia (CFP) tiene antecedentes en la literatura en autores como Millikan (1995), Plantinga (1993). Por otra parte, se pueden encontrar aplicaciones a otras áreas de la epistemología distintas al conocimiento como el trabajo de Graham (2014) sobre el acreditamiento [entitlement] perceptual y el de Bergmann (2006), sobre la justificación (para una revisión de la literatura Cfr. Boyce y Plantinga 2012). De manera informal, la tesis consiste en considerar como conocimiento una creencia que fue producida por un mecanismo cognitivo el cual tiene la función es producir este tipo de creencias.

El tipo de función que tiene en mente el teórico de funciones propias es el de las funciones que se asignan paradigmáticamente en biología. De hecho se puede considerar una aplicación de los desarrollos teóricos de funciones propias en biología para las funciones cognitivas. Podemos pensar en ejemplos obvios, como que la función propia de las alas es volar (en las especies que vuelan); pero también casos donde la función fue producto de un descubrimiento, como la adscripción de la función de bombear sangre para el corazón (Galeno en el siglo II A.C. le adscribió erróneamente la función de proveer calor). Las funciones no son sólo descriptivas, sino normativas aún en los casos en que no se lleva a cabo la función pero ésta existe, i.e. un riñón que no está filtrando la sangre, aún tiene la función de filtrar la sangre.⁶⁷

Dentro de las distintas teorías de función biológica, las que consideran que son funciones propias (también llamadas teleológicas) defienden explícitamente que los rasgos objetivamente poseen dichas funciones: no son atribuidas por los biólogos, sino descubiertas. Desde un punto de vista metafísico, ¿en qué consisten estas funciones que se descubren? La respuesta estándar⁶⁸ es que las

⁶⁷ En las discusiones sobre funciones biológicas se suele considerar como un elemento importante que una teoría también sea capaz de atribuir función en casos de disfunción.

⁶⁸ Una excepción notable es Alvin Plantinga (1993), quien habla de funciones propias apelando al diseño inteligente. Si bien es un tipo de creacionismo que acepta la historia evolutiva, rechaza la selección natural como mecanismo principal para producir rasgos. Por tratarse de una postura *sui generis*, cuando hable de teorías teleológicas o de función propia no tomaré en cuenta esta versión de la tesis.

atribuciones teleológicas son una manera oblicua de hablar acerca de la evolución de dichos rasgos (Ayala, 1977). Así, el lenguaje teleológico de los biólogos puede ser objetivo sin apelar a un elemento teleológico irreducible, como sucedía con las naturalezas [αἰ φῦσεις] de Aristóteles (Cfr. *Física*, 192b20–23). Se trata de una normatividad reducible a hechos categoriales (aunque en ocasiones —dependiendo de la teoría teleológica— históricos). Habiendo realizado estas aclaraciones previas, se puede caracterizar la tesis del conocimiento como función propia del siguiente modo:

CFP: Una creencia verdadera B es conocimiento si y sólo si fue producida por un conjunto de mecanismos de generación de creencias M1, M2... Mn, que tienen la función propia —posiblemente entre otras— de producir creencias del tipo relevante T al que pertenece B.

Hace falta notar algunas cosas sobre dicha definición. En primer lugar, la idea de función no se encuentra explicitada en esta definición. Esto es sumamente relevante debido a que existen diferentes intentos de explicar filosóficamente en qué consiste una función y concretamente una función propia. En segundo lugar, tampoco aparece explicitado a qué se refiere el tipo relevante al que pertenece dicha creencia. La literatura existente se ha concentrado sobre el ambiente, en sentido amplio, entendido como ambiente cognitivo. No obstante, podemos pensar también en dominios, cuando los mecanismos mencionados tengan una especificidad de dominio (ejemplos de dominios pueden ser la aritmética, el ámbito social, etc).

En este capítulo daré una muestra de cómo se puede aplicar la tesis general de CFB para naturalizar el conocimiento con el caso de la detección de depredadores. Posteriormente haré una revisión breve de la literatura sobre funciones biológicas y evaluaré las distintas alternativas para caracterizar una función propia en relación con el problema del conocimiento. Finalmente, mostraré como en

general todas las teorías teleológicas carecen de recursos suficientes para proveer una normatividad epistémica adecuada.

3.2. El comportamiento antipredatorio

Expondré a continuación elementos clave del comportamiento antipredatorio en algunas especies de mamíferos y aves, como ejemplo de un comportamiento complejo que se puede explicar apelando a una noción de CFP.

Un ejemplar de conocimiento es una creencia que fue generada por los mecanismos generadores de creencias que cumplen su función adecuadamente. Por ello, como señala Kornblith (2002), el conocimiento es un fenómeno natural que se limita a las especies a las cuales es adecuado atribuir creencias. Sin entrar en detalle sobre este punto, podemos tomar como paradigma a ciertas especies que poseen una cognición sofisticada, como ciertas aves y mamíferos. Los casos más notables pueden ser grandes simios, delfines, cuervos o loros grises. Pero más allá de estos casos extraordinarios de inteligencia no-humana, resulta más interesante —en aras de la generalidad— extender la atribución de creencias y conocimiento a otras clases de especies como primates, ungulados, roedores, miembros de la clase carnívora y otras especies de aves no tan célebres por sus capacidades cognitivas (ver por ejemplo, el trabajo de Smith y Johnson, 2012 acerca del pollo doméstico). En adelante hablaré de “animales no-humanos” refiriéndome de manera imprecisa al tipo de especies que recaen bajo esta lista incompleta; dejando abierta la cuestión sobre si se puede extender esto a casos menos paradigmáticos (como los pulpos)

A pesar de las analogías, y especialmente las homologías, que puedan existir, los mecanismos cognitivos de cada especie tienen sus particularidades. Un problema consiste en mantener un nivel

de generalidad en esta discusión. Mi propuesta consiste en tomar un fenómeno común que se presente en diversas especies, que son los mecanismos de detección de depredadores.

La atribución de creencias a animales no-humanos es problemática. Tomo aquí al igual que Kornblith (2002), a la etología cognitiva como la disciplina encargada de estudiar intencionalmente a los animales no-humanos. La atribución de creencias realizada en etología pretende ser más que una metáfora, o una manera antropomórfica de describir y/o predecir el comportamiento animal. El problema de tomar a los seres humanos —capaces de utilizar un lenguaje de manera recursiva y por tanto creativa— como paradigma de intencionalidad, es que la tendencia será dar una versión demasiado estricta de los requisitos para poseer actitudes como las creencias. Se puede ver el caso de un hablante competente adulto, que es capaz de concebir distintos estados de cosas que caen bajo una proposición.

Por lo anterior, los contraargumentos para atribuir creencias a individuos que no son usuarios lingüísticos se fundamentan en que “para tener una creencia, hace falta tener una perspectiva del estado de cosas representado” (Cfr. Andrews, 201, p. 58⁶⁹). Consideremos, por motivos de simplificación, que la semántica usual que se utiliza para las creencias es correcta y que el contenido de una creencia es una proposición. Una proposición a su vez está compuesta de términos. Partiendo de que se cumple el principio de composicionalidad o de que hay algo semejante a la composicionalidad, el significado de la proposición es una función del significado de dichos términos y la estructura sintáctica en la que están acomodados. Por tanto —según esta línea de razonamiento— para comprender una proposición hace falta poseer un concepto de cada término involucrado

⁶⁹ Es una exposición de Andrews de la postura, no sostiene la tesis.

Los conceptos humanos tienen un grado de generalidad que no aparece en los ejemplos de conocimiento utilizados en los capítulos anteriores. Recordamos que los frailecillos silbadores estudiados por Ristau (1991) son capaces de identificar depredadores. El trabajo de Ristau muestra que este reconocimiento va más allá de un mecanismo automático o reflejo, ya que el frailecillo reacciona de acuerdo a lo que hace el depredador y orienta su danza de ala rota de modo que pueda alejarlo del nido. No obstante, no hay ninguna evidencia —y no tendría ninguna plausibilidad pensarlo— de que el frailecillo tenga algún concepto de “amenaza” o “depredador” con la generalidad que tiene estos conceptos en hablantes competentes

En el mejor de los casos, este tipo de contenido conceptual se daría en un reducido número de especies de animales no-humanos. Los conceptos o análogos a “conceptos” en animales no-humanos son dependientes a un contexto (Andrews, 2015). Un frailecillo silbador puede identificar depredadores en determinadas situaciones relevantes para ella y sus crías, sin por ello saber lo que es un depredador en abstracto. Mi objetivo no es entrar en detalle acerca del problema de contenido mental en animales, ya que la literatura es amplia (e.g. los trabajos clásicos son Davidson, 1982, Stich, 1979, Dennett, 1996, Fodor, 1975, Bermudez, 2009, etc). Sino dar cuenta de cómo podemos hablar de conocimiento proposicional en animales no-humanos, si su contenido conceptual no posee la riqueza semántica suficiente para conformar proposiciones (del modo usual como se construyen en la literatura filosófica). Ante este tipo de problemas existe una tentación es abandonar el lenguaje intencional y adoptar un naturalismo eliminativista (como el expuesto en el capítulo 1); es decir, redibujar por completo la epistemología. Pero como exponen los capítulos anteriores, el objetivo de esta investigación es intentar construir una epistemología que sea naturalista y no-eliminativista hasta donde sea posible. Para resolver este problema en concreto es posible recurrir a una herramienta que permita hablar de conocimiento proposicional sin necesidad de atribuir esta estructura semántica complicada a las proposiciones. A continuación tomo esta propuesta de Rowlands, 2012, p. 58):

La proposición p rastrea la proposición p^* si la verdad de p garantiza la verdad en virtud del hecho de que hay una conexión asimétrica fiable entre el concepto expresado por el término que ocupa el lugar del sujeto en p y el concepto expresado por el término que ocupa el lugar del sujeto en p^*

De este modo, aunque un animal como el frailecillo no cree la proposición p^* “hay un depredador cerca”, queda abierta la posibilidad de decir que cree una proposición o cuasi proposición —cuyos términos son dependientes a un contexto— que rastrea la proposición p^* . Lo relevante entonces para hablar de conocimiento es mostrar hasta qué punto una cuasi proposición de este tipo puede rastrear la verdad de la proposición relevante. Esto enmarca los términos en los cuales se puede estudiar el conocimiento animal, tomando como estudio de caso el comportamiento antipredatorio.

El comportamiento antipredatorio puede descomponerse en tres grandes etapas: vigilancia, clasificación de predadores y elección de respuesta (Allen y Bekoff, 1997, p. 116). Por supuesto, el comportamiento antipredatorio no siempre es un fenómeno cognitivo complejo, pero aquí me centraré en los casos en los que dicho comportamiento tiene cierta flexibilidad en la capacidad de reconocimiento y clasificación de depredadores (Allen y Bekoff, 1997, p. 117). Es decir, que no es únicamente una respuesta automática.

De estos tres fenómenos el que resulta más interesante para los fines de este capítulo es el de la clasificación de depredadores. La vigilancia es un fenómeno que debemos entender desde un punto de vista de la adquisición de información. Los procesos que constituyen a la vigilancia son interesantes de suyo desde el punto de vista cognitivo y muy particularmente desde el punto de

vista de la distribución del trabajo cognitivo (como en el caso del forraje en grupo). En especies como los antílopes o los chimpancés (Premack, 1986) la vigilancia tiene un carácter especialmente dinámico, ya que la conducta que despliega el individuo está relacionada con el tamaño del grupo del que forma parte. Podemos considerar entonces que el individuo tiene una creencia acerca del tamaño del grupo del que forma parte, y que dicho conocimiento ha sido producido por mecanismos evolutivos cumpliendo con su función. Por tanto dichas creencias acerca del tamaño del grupo constituyen conocimiento

La elección de respuesta también puede llegar a ser un fenómeno cognitivo complejo en diversas especies, en especial cuando la flexibilidad es suficiente para escoger una respuesta no sólo frente al tipo de depredador que es detectado, sino a otros elementos ambientales (como la distancia, el tamaño del depredador y nuevamente el tamaño del grupo del que el individuo forma parte. En adelante me centraré en la parte de clasificación de depredadores, por considerarlo el paso más paradigmático del comportamiento antipredatorio.

Recordemos que la fiabilidad de los mecanismos de detección de depredadores depende de consideraciones evolutivas. Una especie puede tener un mecanismo general, pero en como señalan Allen y Bekoff (1997, p. 121), “bajo condiciones específicas habrá presión selectiva para mayor discriminación entre depredadores.” Dicha detección puede darse, pensando en los extremos, desde mecanismos completamente innatos hasta por aprendizaje. Entre estos dos extremos, se encuentran los mecanismos innatos que deben ser reforzados, especialmente durante la etapa temprana. La creencia adquirida no necesariamente rastrea una proposición acerca de una especie (e.g. “águila”), sino que tiende a ser más general (como la distinción entre amenazas aéreas y terrestres).

Los estímulos obtenidos durante la etapa de vigilancia y a partir de los cuales se generan creencias acerca del tipo de amenaza a la que hay que responder pasan por una etapa de generalización. Como indica Shettleworth (2010, pp. 182- 183), “hay una relación compensatoria [trade off] entre generalizar y discriminar (McLaren 1994) [...] El grado de generalización depende del comportamiento del sistema y de la fuerza de la motivación que subyace a la respuesta” Como se indica anteriormente, el resultado no es algo que propiamente se pueda considerar como un concepto, sino una “discriminación categorial”, (Shettleworth, 2010, p. 191). La discriminación categorial (en general) puede ser con base a características perceptuales (el color, la forma) o funcionales (ser comestible, ser una superficie por la cual se puede caminar). En los organismos capaces de llevar a cabo discriminación categorial, la conducta producto de categorías tiene una complejidad que va más allá de una conducta simplemente detonada por un signo; como el caso de la inmovilidad tónica que presentan como reflejo automático ciertas especies ante movimientos rápidos con el objetivo de evadir un depredador pasando por muertos.

Mi objetivo en este apartado ha sido mostrar cómo se puede buscar fructíferamente una noción de conocimiento, citando literatura tanto de etología cognitiva como de psicología cognitiva. Este ejemplo sirve para mostrar cómo la tesis del conocimiento como función propia provee más flexibilidad para naturalizar el conocimiento. Se trata de una noción más abstracta, que no se identifica reductivamente con una noción de conocimiento particular como en el caso del conocimiento etológico (CE) en la estrategia de Kornblith; pero que permite capturar generalidades en distintas especies, estudiadas por disciplinas diferente. Considero que en esto radica la fortaleza principal de la tesis. En cambio, mostraré en el siguiente apartado, que la manera de asignar funciones apelando a mecanismos de selección, ocasiona problemas importantes que motivarán buscar otra manera de defender la tesis del conocimiento como función propia.

3.3. Las teorías teleológicas de funciones biológicas

La teoría evolutiva respondió al enigma permanente (e.g. Platón, *Filebo* 28c; Hume, DNR 8.9, etc...) acerca de cómo explicar lo que tiene apariencia de finalidad sin un diseñador. El reto contemporáneo es inverso: cómo dar lugar a nociones aparentemente teleológicas —como las funciones— dentro de una teoría que no es teleológica en el viejo sentido. Así, el proyecto de analizar filosóficamente la noción de funciones biológicas más cercano a nuestro tiempo surge a partir del uso extendido que se les da a estas nociones en la biología y ciencias afines (e.g. Ayala, 1977) y a la dificultad de hacer compatibles con el paradigma científico actual —y en particular, de la biología posterior a la teoría sintética de la evolución (TSE)— formulado en términos de explicaciones mecánico-causales. El proyecto ha tomado la forma tanto como un intento de proveer una noción unificadora de función (y que por tanto abarque a las funciones biológicas), como la elaboración de un análisis acerca de las funciones biológicas que no abarca otro tipo de funciones (como las de los artefactos).

Los trabajos preliminares de Ernest Nagel (1961), y Carl Hempel (1965) son el punto de partida del debate contemporáneo acerca de una teoría general funciones (aunque motivados por dar cuenta de las funciones biológicas). Estos trabajos pueden verse como un acercamiento inicial a una literatura amplia que busca definir de manera precisa dicha noción. La naturaleza de estos trabajos no suele ser revisionista, sino que —por el contrario— parte del uso que se le da a las nociones de función biológica en las explicaciones de los científicos practicantes, intentando unificarlas en una noción o en un conjunto pequeño de nociones bien definidas.

Además del antecedente contemporáneo de E. Nagel y Hempel, los trabajos considerados como auténticamente seminales son los de Wright (1973) y Cummins (1975). El análisis de función de Wright es el primero que puede considerarse etiológica, si bien es un tratamiento que encuentra

antecedentes en los trabajos de los propios biólogos (e.g. Mayr, 1961). La idea fundamental es que si la presencia de un mecanismo puede explicarse en términos causales porque realiza una determinada tarea, la tarea puede considerarse como la función de dicho mecanismo. Esta intuición fundamental es la base de teorías etiológicas más completas como las de Millikan (1984) y Neander (1991) y otra gran cantidad de autores (para una lista más amplia ver Buller, 1999), formuladas en términos explícitamente evolutivos.

Por otro lado, el artículo de Cummins "Functional Analysis" (1975) es el trabajo seminal para las teorías sobre funciones que se basan en el análisis funcional, o como lo llama Davies (2003), el enfoque de capacidades. La idea fundamental consiste en mirar a la función como parte de la capacidad de un sistema que presenta un interés teórico, este sistema se puede subdividir en partes, de manera que una función puede entenderse como una contribución causal de una de estas partes hacia un proceso complejo que tiene como resultado la capacidad del sistema que constituye el objeto teórico.

Al igual que los análisis de E. Nagel y Hempel, las propuestas de Wright y Cummins son teorías unificadoras acerca de las funciones. Es decir, aplican tanto a funciones biológicas como de artefactos. No obstante, en trabajos posteriores se han defendido enfoques pluralistas (hay dos nociones independientes de función operando en cada ámbito), unificadores (ambos tipos de funciones pueden entenderse bajo un sólo esquema, como el de diseño Cfr. Kitcher, 1981) o de instanciación; las teorías etiológicas son en realidad una instancia de análisis sistémico sobre poblaciones genéticas (Davies, 2003, pp. 30-33).

Los otros dos enfoques —también seleccionistas— notorios son la propuesta de Kitcher (1981) anteriormente mencionada, y las teorías propensionistas. En el primer caso, hay una noción de diseño a partir de la capacidad de la selección natural de producir adaptaciones. Un agente

intencional al tener un diseño en mente para un artefacto determina su función de este modo — una tesis controvertida en la literatura sobre artefactos, pero plausible—, la selección natural cuando funciona de manera adaptativa explora diseños a través de un espacio de posibilidades (aquellas que proveen mayor aptitud). La idea principal de Kitcher es englobar ambas cosas bajo una noción más abstracta de diseño.

Dentro de las teorías anteriormente descritas, en este capítulo tomaré a las teorías seleccionistas como paradigma de función propia para sostener CFP. En particular, me centraré en las teorías etiológicas, por las teorías seleccionistas de función mayormente defendidas en la literatura. Asimismo, las teorías etiológicas son las que se han tomado en los pocos trabajos de epistemología que sostienen CFP, siendo el texto de Ruth Millikan (1995) "Naturalist reflections on knowledge" el antecedente más importante de una epistemología teleológica. A continuación daré una breve exposición y clasificación de las teorías etiológicas.

Buller (1999, p. 281) clasifica las teorías etiológicas en dos grandes grupos: fuertes y débiles, haciendo notar que algunos de sus defensores no suelen tener presente esta distinción. Ambas teorías consideran que las estructuras biológicas tienen funciones propias, dichas funciones son determinadas por la historia de dicho rasgo; incluyendo a los genes causalmente responsables de su formación. Los genes tienen un papel especial en las teorías teleológicas debido a que los rasgos no producen otros rasgos (un corazón no produce corazones).

Para mostrar cómo los genes forman una parte central de las teorías etiológicas ilustraré esto con el trabajo de Millikan. Los genes responsables de un rasgo son "familias reproductivamente establecidas de primer orden", es decir "un conjunto de entidades que tienen los mismos caracteres reproductivamente establecidos o similares, derivados por reproducciones repetidas del mismo modelo o modelos" (Millikan, 2001, p. 23). Cuando la función propia de las familias

reproductivamente establecidas de primer orden es producir determinados ítems, estos ítems forman junto con esta familia de primer orden, una familia reproductivamente establecidas de orden superior. A esta familia de orden superior es a la que se le atribuye una función propia. En términos informales, el corazón tiene la función de bombear sangre dado que pertenece a una familia reproductivamente establecida (el corazón y los genes responsables para producir corazones) cuya función propia es bombear sangre.

Haciendo de lado estos detalles que forman parte de una explicación más granular, la presencia de un rasgo se explica debido a su contribución a la aptitud de los individuos que lo poseían. Las teorías etiológicas débiles pueden formularse del siguiente modo (Buller, 1999, p. 284)

Un ejemplar presente de un rasgo T en un organismo O tiene la función de producir un efecto de tipo E sólo si los ejemplares pasados de T contribuyeron a la aptitud de los ancestros de O debido a que producían E, y por tanto contribuyeron causalmente a la reproducción del rasgo T en el linaje de O.

Las teorías etiológicas fuertes tienen un requisito adicional: T debe ser haber sido seleccionado entre otros ítems alternativos porque (debido a que) contribuía a la aptitud de los ancestros de O (Buller, 1999, p. 283; Godfrey-Smith, 1994, p. 359). Lo anterior implica la existencia de alternativas en el acervo genético de una misma población que enfrenta el mismo ambiente selectivo (Brandon, 1994). En palabras de Buller (1999, p. 284)

Si la variación de T se encontraba distribuida por grupos [...] y cada grupo se enfrentaba a distintas demandas ambientes, las diferencias en supervivencia y reproducción no fueron debido a diferencias en aptitud de los miembros de dichos grupos, sino a las diferencias en los ambientes que enfrentaron.

En términos más generales, mientras que las teorías débiles determinan las funciones con base a la historia evolutiva, las teorías fuertes se limitan a la historia selectiva. Por ello, las teorías débiles pueden asignar funciones a rasgos que o bien no fueron seleccionados sobre otras alternativas (aunque enfrentaron presiones selectivas), o bien que fueron seleccionados en una población que no compartía el mismo ambiente selectivo, o bien —en el caso más débil— que se fijaron por otros mecanismos como la deriva genética.

¿Qué ventajas o desventajas pueden tener estas teorías para utilizarse de manera normativa para juzgar las funciones cognitivas (y de manera concreta, la generación de creencias)? En adelante, mostraré que tanto la versión fuerte de las teorías etiológicas como la versión débil son inapropiadas —por razones distintas— para utilizarse en una teoría del conocimiento.

Por una parte, las teorías fuertes, al ser más restrictivas, dejan sin adscribir funciones a una cantidad de mecanismos a los que las teorías débiles sí podrían adscribirles una función. Por otra parte, las teorías débiles tienen la desventaja de que pueden sobredeterminar las funciones. En este momento no llevaré a cabo una distinción precisa entre ambas familias de teorías, ni tampoco entre teorías particulares; no obstante, si los problemas más generales de las teorías etiológicas de función pueden resolverse será necesario adentrarse en estos detalles para ver una propuesta sustancial de una naturalización del conocimiento.

Un elemento importante de cualquier teoría, sea o no etiológica, consiste en la posibilidad de atribuir función aún en los casos en que hay disfunción. Por ejemplo, el ojo de un individuo que no transmite información al nervio óptico no está realizando su función (la vista), pero ésta no deja de ser su función. Por tanto, para atribuir disfunción —lo cual es un tipo de evaluación— es necesario atribuir función.

El problema en este caso, radica en distinguir entre un caso de auténtica disfunción y un caso donde hay una estructura que es vestigial (es decir, que ya no realiza su función original en dicha especie). Como señala Davies (2003, p. 196), la propuesta que aparece en Griffiths (1992) y otros defensores de las teorías seleccionistas (entre las cuales se cuentan las etiológicas), depende de que haya pasado suficiente tiempo para que en la población donde se encuentra un determinado rasgo, dicho rasgo haya dejado de contribuir a la aptitud genética llevando a cabo su función. Por ejemplo, los músculos de la oreja están presentes pero no son funcionales en los seres humanos y chimpancés, pero dada la gran cantidad de tiempo que tienen sin contribuir a la aptitud, no se consideran una disfunción.

Uno de los aspectos más importantes de buscar una teoría adecuada sobre funciones tiene que ver con el aspecto normativo: de qué modo se puede evaluar si un mecanismo está cumpliendo con su función correctamente. La motivación principal es dar cuenta de qué modo dicha normatividad puede ser compatible —y en última instancia fundamentar— una normatividad epistémica en el caso particular del conocimiento.

Para ello, será útil mirar a los mecanismos generadores de creencias bajo la óptica de funciones biológicas o al menos naturales. Como se verá más adelante, esto no conlleva un compromiso con las teorías etiológicas. Las teorías de análisis sistémico también son una alternativa para elaborar una naturalización del conocimiento. No obstante, para hablar de naturalización de manera acorde con los criterios establecidos en el primer capítulo, será necesario mostrar que los análisis de funciones en el espíritu propuesto por Cummins no son arbitrarios ni dependen exclusivamente de los intereses de quien realiza la atribución. En el siguiente capítulo mostraré cómo dichas teorías son compatibles con el naturalismo; y por otra parte, cómo permiten elaborar una epistemología de la virtud naturalizada

3.4. La evaluación cognitiva con funciones biológicas

Antes de considerar de qué manera las distintas teorías de funciones biológicas se acomodan dentro de una epistemología naturalista, es necesario señalar algunas cosas sobre el papel normativo que poseen las evaluaciones de función. La normatividad del conocimiento radica en que por una parte, es un objetivo deseable saber que P, más aún, es preferible saber que P que meramente creer verdaderamente que P. Por otra parte, las evaluaciones de función tienen su propia normatividad. Las nociones de buen y mal funcionamiento evalúan si un mecanismo realiza lo que debe hacer del modo correcto.

El objetivo entonces es explorar hasta qué punto la normatividad epistémica puede fundamentarse en esta segunda normatividad, sin apelar a los elementos fiabilistas presentes en Kornblith. El problema tiene —a grandes rasgos— tres aspectos. En primer lugar, es necesario determinar si es posible explicar la diferencia entre una creer verdaderamente que P y saber que P para un conjunto creencias de tipo básico (como las creencias perceptuales, o las creencias inferenciales sencillas).

La tesis del conocimiento como función puede formularse dentro de la epistemología tradicional, como creencia verdadera con un requerimiento añadido que distingue al conocimiento de otras creencias verdaderas. En el caso de Millikan (1995, p. 241)⁷⁰, esta tesis se formula como “creencia verdadera adquirida de modo Normal.” El uso estipulativo del término “Normal” —en adelante escribo *normal*, con itálicas—, incluye principalmente a las funciones propias de los mecanismos biológicos básicos, pero también incluye elementos externos como ayudas perceptuales (anteojos, aparatos de sordera) y teorías científicas. A pesar de que estos mecanismos formadores de creencias son externos, el modo como sopesamos distintas disposiciones a creer que P —por

⁷⁰ Millikan presenta su versión como una variante de fiabilismo, pero el énfasis de su propuesta está en los elementos teleológicos.

ejemplo, al contrastar nuestra impresión inocente de que la Tierra es plana con el consenso científico— es a través de mecanismos de evaluación de creencias que son en última instancia naturales. Por ello, Millikan (1995, p. 264) considera que su definición de conocimiento engloba a un sólo fenómeno y que éste es un fenómeno natural.

El trabajo de Millikan ilustra una estrategia razonable que podría seguir una naturalización del conocimiento. Existe un cierto parecido con las teorías fundacionistas del conocimiento, donde hay una base de creencias que se extiende por medios inferenciales (no necesariamente deductivos, Cfr. Audi 1993). No obstante, hay una diferencia importante. El fundacionismo epistemológico, al igual que su contraparte —el coherentismo—, es un análisis de las creencias de un agente en un momento dado (i.e. de una rebanada temporal); en otras palabras, es un análisis estático.⁷¹ En la explicación de Millikan, las creencias básicas juegan un papel histórico-causal. Las creencias producidas por medios que no son parte de nuestro aparato cognitivo (como las teorías científicas) son normales, porque a su vez fueron generadas a partir de creencias normales y son mantenidas de modo normal por nuestros mecanismos inferenciales en combinación con estos elementos externos.⁷²

Considero que una naturalización exitosa del conocimiento (CN) utilizando las herramientas de CFP, idealmente debe cumplir con los siguientes puntos:

- A) Los mecanismos cognitivos responsables de generar creencias tienen funciones propias.
- B) Las creencias verdaderas generadas por mecanismos cognitivos llevando a cabo su función propia cumplen con los desiderata del conocimiento.

⁷¹ En cuanto a que se refiere al apoyo que recibe una creencia de otras en un momento dado, Cfr. Pollock y Cruz (1999). En este mismo sentido, la teoría fiabilista de Goldman también tiene un componente histórico-causal.

⁷² Desde un punto de vista superficial esto da la apariencia de circularidad, pero podría ser formulado adecuadamente de manera recursiva.

- C) El conocimiento proposicional humano se agota en las creencias generadas como indica b.
- D) Podemos conocer la función propia de los mecanismos cognitivos.
- E) Podemos atribuir conocimiento aún sin conocer la función propia de los mecanismos cognitivos.

Los puntos A y B son imprescindibles para llevar a cabo el proyecto de naturalización. A favor de A, se puede decir que *prima facie* podemos asumir que es una consecuencia de la tesis de que las funciones biológicas son funciones propias. Es decir, mientras no haya razones en contra, debemos tratar al aparato cognitivo como un rasgo biológico más. Considero que es un buen principio para investigar las funciones cognitivas, y que si las teorías de función propia son las correctas para asignar funciones biológicas también deben serlo para asignar funciones cognitivas. En el siguiente apartado problematizaré A, sugiriendo que hay razones para cancelar ese *prima facie*; es decir, que determinadas funciones cognitivas tienen dificultad para asignarse con las teorías de función propia. Mi punto de vista es que esto no es una razón suficiente para hacer una teoría de función ad hoc para las funciones cognitivas, puesto que otras teorías de función más cercanas al análisis de capacidades sistémicas de Cummins pueden explicar tanto las funciones biológicas en general como las funciones cognitivas. No obstante, no proseguiré esa línea de crítica, sino que la fuerza de mi rechazo a CFP se centrará sobre B, mostrando como el conocimiento resultante de dicha teoría no incluye los fenómenos adecuados desde los mínimos que buscamos para un punto de vista epistémico-normativo; es decir, no desde una noción preteórica fuerte o de un análisis procedente de la epistemología tradicional, sino de los desiderata básicos indicados en el capítulo 1.

El punto C, a diferencia de A y B, no es un requisito indispensable para naturalizar el conocimiento como sugiere CFP, pero sí es deseable —al menos en una versión moderada. En una versión moderada, C indica que además de que las creencias verdaderas que cumplen con A y B son

conocimiento, no hay o bien un conocimiento distinto en tipo (por ejemplo, que requiera evidencia accesible al agente) o un conocimiento de un nivel más alto (que requiera una elaboración reflexiva de los conocimientos que cumplan con A y B, pero que a su vez no sea producto de A y B⁷³, como considero que sería el caso de la división de Sosa de conocimiento humano y animal). En una versión extrema, CFP engloba además cualquier tipo de conocimiento, incluyendo el conocimiento científico, las matemáticas avanzadas, etc. No es necesario sostener la versión extrema de la tesis, dado que uno puede decir que en estos casos hay algo *más* que conocimiento. Por ejemplo, a nivel del conocimiento tanto el lego que se entera por una buena fuente que los números primos son infinitos como el matemático que está familiarizado y entiende alguna de las pruebas disponibles (como la de Euclides), saben por igual que los números primos son infinitos. El matemático tiene una mejor situación epistémica respecto a la misma proposición, pero esa mejor situación se puede explicar en otros términos (como la justificación o la comprensión). Es decir, la situación epistémica del matemático va más allá del conocimiento.

No problematizaré D, porque considero que es una hipótesis plausible. Cabe mencionar que hay casos minoritarios pero notables de científicos cognitivos que son escépticos acerca de identificar la función adaptativa de los rasgos cognitivos, como Fodor y Chomsky (Cfr. Dennett 1996, cap. 13). No obstante, sus críticas han tenido poco eco entre la comunidad científica.

Tampoco desarrollaré E, debido a que las atribuciones epistémicas son un tema que merece ser tratado de manera independiente, pero cabe mencionar que las críticas de Braddon-Mitchell y Jackson (2006) hacia las teorías teleológicas de contenido (teleosemánticas) se pueden aplicar igualmente a las atribuciones de conocimiento. De acuerdo con la objeción de opacidad referencial

⁷³ Esta aclaración es importante. Existen mecanismos de metacognición y monitoreo que en ocasiones pueden ser centrales para adquirir algunos tipos de conocimiento. Estos mecanismos están llevando a cabo su función propia, por lo cual no es un contraejemplo para CFP. No obstante, si se lee la reflexión como un requisito internista más allá de estos mecanismos de control, sí hay incompatibilidad con CFP.

de Jackson, dado que las teorías teleosemánticas determinan el contenido por medio de la historia evolutiva, quienes no conocen la historia evolutiva (¡o siquiera la teoría de la evolución!) no pueden hacer atribuciones de contenido. Las teorías de rol funcional no tienen este problema, debido a que las atribuciones folk de contenido —de acuerdo con Jackson— son funcionales (de manera imprecisa). La teoría del conocimiento que defenderé en el siguiente capítulo basada en capacidades sistémicas, de manera análoga a las teorías de contenido de rol funcional, es compatible con las atribuciones ordinarias de conocimiento.⁷⁴

3.5. Las funciones seleccionadas no son una condición suficiente

El objetivo de este apartado es establecer que las funciones seleccionadas no pueden proveer condiciones suficientes para el conocimiento. Cabe notar que en el marco metodológico de esta discusión no se pretende llevar a cabo un análisis en sentido estricto, por lo cual parecería que no hace falta sostener algo tan fuerte como que las tesis del naturalista ontológico determinen condiciones necesarias y suficientes que correspondan con nuestras atribuciones ordinarias de conocimiento. No obstante, mi intención aquí no es mostrar contraejemplos extravagantes, en mundos posibles o escenarios inusuales para refutar un análisis conceptual. Mi objetivo es mostrar que el fenómeno del conocimiento (no nuestro concepto de él) se aparta de maneras importantes de lo que las teorías teleológicas en sus distintos sabores y variedades pueden reconocer como conocimiento, no a través de lo que *llamaríamos* conocimiento en un experimento mental dentro de uno de estos escenarios bizarros, sino de que lo que regularmente llamamos conocimiento para cumplir con los fines evaluativos y normativos del término. Recordando la estrategia de

⁷⁴ Moisés Macías-Bustos me ha sugerido en comunicación personal que la idea de Lewis (1983) de que hay un magnetismo de referencia —i.e. que la referencia de los términos es la que respeta las coyunturas de la naturaleza— se puede usar para resolver este problema. No estoy seguro de que pueda aplicarse para resolver el problema análogo en el caso de las atribuciones de conocimiento.

acomodación descrita en el capítulo 1, buscaré mostrar que la noción resultante no puede llevar a cabo el trabajo teórico de una noción adecuada de conocimiento.

Comenzaré por el aspecto de la suficiencia de las funciones seleccionadas para el conocimiento. La propuesta de Kornblith combinaba tanto elementos teleológicos como fiabilistas. En el capítulo anterior fue posible mostrar con algunos ejemplos cómo el fiabilismo dejaba de atribuir conocimiento en casos de sistemas que —a pesar de que realizan su función correctamente— no la realizan de manera fiable (entendiendo la fiabilidad como una tasa de éxito sobre fracasos). No obstante, a pesar de los problemas señalados en el capítulo anterior a FE, la fiabilidad es un elemento que contribuye a cumplir con los desiderata más generales del conocimiento establecidos en el primer capítulo, particularmente la condición de seguridad. El hecho de que un mecanismo cognitivo genere una creencia verdadera del tipo relevante, realizando la función para la que dicho mecanismo fue seleccionado, no descarta que dicha creencia haya sido obtenida por suerte epistémica por dicho individuo.

En resumen, la selección natural no puede explicar por qué los mecanismos seleccionados son una condición suficiente para proveer la normatividad epistémica que buscamos en el conocimiento. A pesar de que es trivialmente cierto que muchas de las modificaciones evolutivas de una especie dan como resultado estructuras fenotípicas que constituyen una respuesta al medio ambiente, hablar de adaptaciones no es tan sencillo como parece. La adaptación no es uno de los mecanismos fundamentales de la teoría sintética de la evolución (TSE), sino la aptitud. La parte central de la TSE puede definirse del siguiente modo:

a) Variación hereditaria: La variación puede ser por recombinación de genes existentes en el acervo genético de una población o por surgimiento de nuevos genes por mutación.

b) Selección natural: La selección natural produce cambios permanentes en el acervo genético. Los

alelos con mayor aptitud son aquellos que se fijan. La aptitud se define en base al éxito que tiene un gen en reproducirse en un ambiente determinado. Realizar una definición no-circular de aptitud es algo que sólo se ha logrado con algún grado parcial de éxito, pero no será necesario aquí entrar en el problema de dar una definición precisa. En este momento no es necesario entrar en problemas relativos a los candidatos para la unidad de selección (individuo, gen, célula, etc), y mecanismos complementarios de selección como la deriva génica (más adelante me concentraré en el gen como unidad de selección).

El adaptacionismo puede entenderse como la tesis que indica que “Los cambios evolutivos en el acervo genético de una población frecuentemente ocurren en la dirección de mayor adaptación” (Ayala, 1970, p. 6), donde adaptación se entiende aquí como la capacidad de una población a enfrentarse mejor al medio ambiente para sobrevivir y reproducirse.

Es claro que la dirección de la selección natural hacia mayor aptitud no siempre produce adaptaciones, pues en ocasiones se fijan alelos que son dañinos para la especie, lo cual además es una evidencia a favor para tomar al gen como unidad de selección (Dawkins, 1976). No obstante, la afirmación únicamente indica que las adaptaciones ocurren con frecuencia.

¿Cómo se conectan la aptitud y las adaptaciones? Dennett (1996) provee una buena imagen para ilustrar la explicación ortodoxa de cómo sucede esto: la variación hereditaria explora un espacio de posibilidades de manera ciega hasta que una de ellas produce mayor aptitud, lo cual hace fijarse a la variación. Mediante este proceso algorítmico, si el medio ambiente permanece igual, el acervo genético llega a un punto en el cual el genotipo no puede tener mayor aptitud.

A partir de lo anterior, resulta tentador caer en lo que Gould y Lewontin (1979) llamaron un programa de investigación adaptacionista, el cual requiere de las siguientes tesis:

- 1- Es posible identificar e individuar rasgos en los organismos. Existen algunos genes que

corresponden a estos rasgos.

2- Por tanto la selección natural al operar sobre una parte del genotipo está operando sobre el rasgo.

3- Mayor aptitud tiende a producir mayor adaptación.

4- Por tanto, si podemos identificar un rasgo que parece ser una adaptación al medio ambiente, fue seleccionado *para* ello.

Cabe notar que nada en la TSE, ni siquiera en la ilustración citada de Dennett, sugiere lo anterior. Contra esta visión demasiado confiada en el adaptacionismo, cabe recordar que los genes comúnmente realizan múltiples funciones y no corresponden únicamente a rasgos. Algunas de estas funciones tienen que ver simplemente con activación de otros genes o con el desarrollo (Noble, 2008). Asimismo, no todos los rasgos fenotípicos surgen como una adaptación, sino que algunos son *exaptaciones* (Gould y Lewontin, 1979), rasgos que surgieron debido a que contribuían en una función y son cooptados para realizar otra función. La imagen se complica aún más tomando en cuenta mecanismos como aquellos estudiados por la teoría de juegos evolutiva o el papel de la deriva génica.

A lo anterior podemos agregar un problema de destacamiento causal. Lo que sufre un cambio evolutivo es el acervo genético de una población, como un todo. Para algunos rasgos cognitivos (en especial si son modulares) es posible examinarlos y determinar si la función que llevan a cabo es adaptativa, pero esto es sólo una parte de la historia. Toda estructura compleja tiene un costo para el organismo en términos de desarrollo y mantenimiento.

Para un mecanismo auténticamente adaptativo es verdadero el contrafáctico de que si no hubiera ejecutado la función que ejecuta de manera ventajosa para los organismos que la poseen, los alelos necesarios para la construcción de dicho mecanismo no se hubieran llevado a cabo

progresivamente en la historia evolutiva de la especie. Pero también son verdad otros contrafacticos: si los requerimientos calóricos de dicho mecanismo (y sus ancestros en cada momento) no fueran satisfacibles por el medio ambiente tampoco existiría dicho mecanismo en la especie; si no hubiera un equilibrio de Nash que permitiera que dicho rasgo (y sus ancestros en cada momento) aumentaran la aptitud de los organismos poseedores tampoco se habría fijado el rasgo, etc.

3.6. Las funciones seleccionadas no son una condición necesaria

La selección natural, entendida como un proceso de selección genética, es insuficiente para asignar todas las funciones que necesitamos para realizar evaluaciones epistémicas. Esto se ve de manera particular en el caso humano, donde factores como la ontogénesis y la plasticidad neuronal juegan un papel importante en la constitución de las funciones cognitivas. Tengo aquí en mente la literatura neuroconstructivista acerca de desarrollo cognitivo humano que ejemplifica el trabajo de Sejnowski, la cual considero que provee buenas razones para moderar el nativismo acerca de las funciones cognitivas y/o cuestionar la modularidad de estas capacidades (Cfr. Karmiloff-Smith, 1994). Los mecanismos de representación en un adulto dependen de un proceso de maduración neuronal, no sólo en el sentido de que éste es una condición necesaria para activarlo sino que el ambiente determina el tipo de estructuras representacionales que se forman en la corteza prefrontal. Esto se puede expresar de manera sintética como dos tesis, de acuerdo con Quartz y Sejnowski (1997):

- 1) El desarrollo es un incremento progresivo de las estructuras que subyacen a la complejidad representacional.

2) Este incremento depende de una interacción con un ambiente estructurado para dirigir el desarrollo.

Para que el argumento tenga peso, no es necesario aceptar un neuroconstructivismo fuerte que rechace la modularidad o el innatismo en general. Basta con sostener que algunas estructuras representacionales relevantes para el conocimiento son el resultado del desarrollo en los términos descritos por 1 y 2, una tesis que tomaré aquí como plausible. ¿Cómo puede el defensor de CFP responder a este reto? Se puede tomar lo que dice Millikan (1988) acerca del aprendizaje para aplicarlo al desarrollo. Millikan considera que una conducta producto del aprendizaje puede tener una función propia heredada del mecanismo responsable del aprendizaje si dicho mecanismo a su vez tiene como función propia el aprendizaje de dicha conducta. No obstante, la perspectiva neuroconstructivista va más allá de mecanismos innatos para adquirir creencias de determinado tipo, sino que pueden ser entendidos como la constitución de mecanismos nuevos. Ahora bien, podríamos extender un paso más allá lo que dice Millikan y mirar la situación así: hay mecanismos cuya función es producir (al modo como describen los neuroconstructivistas) otros mecanismos, los cuales heredan la función propia de manera análoga a como sucede en el caso del aprendizaje. Sin embargo esto nos obligaría a excluir casos en los que utilizamos mecanismos cuya función adaptativa es aprender en un contexto c_1 para adquirir creencias verdaderas en un contexto completamente diferente c_2 que no fue parte de la historia selectiva; donde a pesar de que c_2 no fue seleccionado, es un contexto donde dicho mecanismo funciona cumpliendo con los desiderata del conocimiento.

Existe otra alternativa, reelaborando la noción de función propia. Si tomamos a las teorías teleológicas como algo que habla únicamente acerca de las funciones seleccionadas en el sentido evolutivo usual, tendremos el resultado de que hay estructuras representacionales que producen creencias verdaderas que cumplen con los desiderata que quisiéramos para el conocimiento, cuyo

éxito epistémico no es producto de la suerte, pero que no cumplen con CFP. Alguien podría defender una variante de la tesis, CFP', donde la función propia no es exclusiva de la selección natural en este sentido restringido, sino de la selección en un sentido más general. Las presiones ambientales para el aprendizaje descritas en trabajos neuroconstructivistas como el anteriormente citado asignan funciones propias, incluso en casos donde éstas no son funciones biológicas (como en el caso de la evolución cultural). De hecho, entendiendo la función propia de este modo, sus alcances rebasan incluso las funciones del aparato cognitivo humano y pueden apelar a las funciones de sistemas externos, como las teorías científicas. Notemos que el defensor de CFP' incluso podría apelar a explicaciones "evolutivas" de la adecuación empírica de las teorías científicas (originalmente formuladas para rechazar el realismo científico como única explicación del éxito de la ciencia). Como indica van Fraassen en *The Scientific Image* (1980, p. 40):

El éxito de las teorías científicas actuales no es un ningún milagro [...] puesto que cualquier teoría científica nace en medio de una feroz competencia, una jungla sangrienta de garras y colmillos. Sólo las teorías exitosas sobreviven —aquellas que de hecho se han enganchado a las regularidades que de hecho existen en la naturaleza.

Lo anterior sería similar a la epistemología evolucionista de Karl Popper (1972) y también de Donald Campbell (1974, 1997), este último no sólo habla del desarrollo de la ciencia sino que trata explícitamente la ontogénesis como un proceso evolutivo; (aunque sería una epistemología evolucionista a nivel de teorías y no de mecanismos) Considero que CFP' es una línea de respuesta viable, aunque ya no se ajusta a la definición de CFP dada en este capítulo ni posee el mismo espíritu naturalista. Hay que tener en mente que la intención original era asimilar el conocimiento a las funciones cognitivas entendidas como funciones biológicas. Esto no es un punto en contra de la teoría del conocimiento propuesta por Millikan, de hecho considero que su capacidad de unificar distintos tipos de conocimiento con orígenes diversos (innato/no-innato, natural/cultural, etc); aunque una vez que se tiene en mente la amplitud de la teoría los resultados pueden ser

decepcionantes para quien esperaba una vindicación inequívoca del naturalismo ontológico para el conocimiento.

Por otra parte, considero que existen razones para rechazar CFP que no se pueden salvar apelando a una interpretación amplia de la noción de función propia; es decir, argumentos que aplican tanto a CFP como a CFP'. Quiero mostrar a continuación que es posible atribuir conocimiento en casos donde los mecanismos son maladaptativos. Centrándonos en el caso de los agentes humanos, existen diversas condiciones en individuos neurodivergentes que crean una diferencia relevante en sujetos neurotípicos. Por razones clínicas, la literatura acerca de los casos donde el razonamiento es deficiente es la más abundante. Destaca aquí el fenómeno de los *delirios*, definidos en el DSM-V (American Psychiatric Association, 2013) del siguiente modo:

Una creencia falsa basada en una inferencia incorrecta sobre la realidad que es firmemente sostenida a pesar de lo que casi todo el resto de las personas creen, y a pesar de aquello que pueda constituir una prueba incontrovertible o una evidencia obvia de lo contrario.

Como se puede ver en la definición, lo más determinante para que una creencia irracional constituya un delirio, es su persistencia. Es decir, como una omisión en el proceso de revisión de creencias, que podemos entender como un tipo de fallo inferencial, a grandes rasgos. Un ejemplo clásico es el delirio de Capgras, donde el sujeto está convencido de que sus familiares más cercanos son personas que fingen ser ellos o intentan reemplazarlos. La creencia es impenetrable a cualquier evidencia que se le presente al sujeto.

Al pensar en las consecuencias que tiene un fallo cognitivo como el delirio de Capgras para la vida de un sujeto, podemos considerarla como una inferencia maladaptativa, desde el punto de vista del organismo. Lo anterior no significa que los genes parcialmente responsables de producir

mecanismos maladaptativos sean de suyo maladaptativos. Por ejemplo, Nesse y Williams (1997) hacen notar estudios que sugieren que los familiares inmediatos de personas con esquizofrenia tienden a ser más exitosos que la población general. Esto deja la puerta abierta a la posibilidad de que los genes que producen la esquizofrenia tengan un valor adaptativo en general, y que la minoría de los individuos que desarrolla la condición sea simplemente un costo colateral (Nesse, 2004). Aún si esta hipótesis fuera correcta, para el individuo que presenta esquizofrenia, sus mecanismos cognitivos claramente no tienen un valor adaptativo (como se puede ver en el hecho de que su esperanza de vida es 12 años menor que el resto de la población).

Por otra parte, tenemos ejemplos de condiciones que podríamos llamar en general maladaptativas, pero que hacen que los individuos que las tienen ejecuten mejor en algunos dominios, como el caso del razonamiento espacial abstracto en autistas funcionales (Stevenson y Gernbacher, 2013). El evolucionista —que busca defender una visión ecológica de la racionalidad basada en la evolución— puede responder a este caso especulando que el autismo no es el responsable de suyo de esta mejoría performativa, sino que se trata ya sea de una compensación en el desarrollo por otras competencias deficientes (en este caso, lingüísticas y sociales), o de que simplemente los mecanismos de razonamiento espacial pueden ejecutar mejor cuando no hay otros presentes (por ejemplo, por tener menos competencia para utilizar recursos de memoria de trabajo y atención).

Por lo anterior, será más conveniente un ejemplo donde una inferencia racional sea producida por los mecanismos divergentes mismos. Para ello sugiero revisar el caso del trastorno de personalidad limítrofe (TPL). Aunque no todos los datos son uniformes, se puede ver en estudios que los individuos con TPL son mejores realizando inferencias sobre estados emocionales negativos a partir de rostros (Scott et al., 2011), algo que se puede considerar como un área de destacamiento entre los sujetos que presentan esta condición (y que explica las descripciones anecdóticas de que los sujetos con DPL pueden “leer la mente”). A pesar de que los sujetos con TPL presentan otros

sesgos cognitivos que pueden llevar a inferencias incorrectas, tal como la tendencia a cometer falacias de falsa dicotomía (pensamiento en “blanco y negro”), podría argüirse que esta capacidad no es de suyo maladaptativa. No obstante, una capacidad destacada no necesariamente implica mayor adaptación. Por ejemplo la hipertimesia (una condición no relacionada con el TPL), consiste en una memoria semántica y/o autobiográfica con capacidades extremas; no necesariamente es mejor que una capacidad de memoria que filtra las cosas irrelevantes. Del mismo modo, un conjunto de mecanismos que detecten “excesivamente” bien las emociones negativas pueden producir un exceso de ansiedad y desconfianza en los sujetos con TPL.

3.7 Conclusión del capítulo

El desarrollo de este capítulo muestra que las teorías que defienden al conocimiento como función propia (CFP) o teleológicas evitan problemas importantes a los que se enfrenta la teoría de Kornblith, mismos que fueron expuestos en el capítulo 2. Además, las teorías teleológicas ofrecen recursos importantes para llevar a cabo evaluaciones epistémicas sin apelar a la fiabilidad. No obstante, CFP hereda los problemas que tienen las teorías de función propia acerca de funciones biológicas. En particular los problemas de las teorías etiológicas, que en su interpretación débil subdeterminan las funciones y en su interpretación fuerte imponen estándares demasiado fuertes para asignar funciones.

También consideré el reto que procede de las teorías neurconstructivistas. Hay que tener en cuenta que se trata de un tema abierto y lleno de controversias (Cfr. Samuels, 2002), pero aun así considero que el defensor de CFP debe darle importancia a las objeciones que se pueden desprender del neurconstructivismo. Si lo que mostré en este capítulo es correcto, el defensor de CFP se enfrenta al dilema de o bien no poder asignar funciones a los mecanismos producidos de este modo, o bien

cambiar la noción de función propia por otra cuyos resultados para la naturalización del conocimiento son decepcionantes.

Finalmente, expuse el caso de conocimiento producto de mecanismos maladaptativos como la razón más fuerte para buscar una alternativa a CFP. Mi propuesta en el siguiente capítulo será explorar la tesis de que el conocimiento se puede naturalizar apelando a las funciones cognitivas, pero no con las teorías teleológicas de función sino con una teoría de función basada en capacidades sistémicas. Con ello no pretendo borrar la importancia de la evolución en una epistemología naturalizada. La historia selectiva nos puede mostrar por qué surgieron ciertos mecanismos cognitivos y por qué son virtuosos (i.e. por qué tienen las capacidades que tienen), pero no qué es lo que los hace virtuosos. Es decir, la historia evolutiva explica desde un punto de vista causal cómo es que podemos tener conocimiento, pero no explica qué es lo que constituye al conocimiento.

CAPÍTULO 4: EL CONOCIMIENTO COMO FUNCIÓN IDENTIFICADA DE CAPACIDADES SISTÉMICAS

4.1. Introducción

El capítulo anterior muestra los problemas que tienen las teorías seleccionistas, y particularmente las etiológicas, para atribuir funciones biológicas. Las funciones cognitivas heredan estos problemas y además presentan otros impedimentos adicionales. Como se puede ver en el capítulo 3, además de los casos de subdeterminación de funciones que impiden llevar a cabo atribuciones de conocimiento (casos donde no podemos saber si una creencia verdadera es conocimiento por no poder evaluar la función de los sistemas que la produjeron), tenemos casos donde las teorías de conocimiento como función propia (CFP) atribuyen conocimiento inapropiadamente o niegan que exista conocimiento contraviniendo los desiderata básicos que expuse en el capítulo 1. Una manera de expresar este fracaso, es que el funcionamiento adecuado del aparato cognitivo no corresponde con las virtudes epistémicas que esperaríamos encontrar en una teoría del conocimiento (Cfr. Graham, 2014 para una visión en contra de esta conclusión).

La vía de solución que exploro a continuación consiste en mantener la tesis de que las evaluaciones de funciones cognitivas pueden proveer una normatividad epistémica adecuada y coincidir con dichas virtudes, pero dejando de lado el compromiso con una teoría teleológica de funciones. Como señalé en el capítulo 3, las teorías de análisis funcional derivadas del trabajo seminal de Cummins (1975) representan la alternativa más prominente en la literatura para elaborar el concepto de función biológica.

Defenderé una teoría de evaluación de las funciones cognitivas con una teoría de capacidades sistémicas muy cercana a la de Cummins, que es la expuesta por Davies en *Norms of Nature* (2003). Desde el punto de vista más general, lo que caracteriza a la teoría de Cummins es que las funciones

se constituyen por su rol causal, por lo cual suele recibir este nombre en la literatura. No obstante, en sus trabajos sobre el tema Cummins expone su teoría utilizando un lenguaje de capacidades y enfatiza la descomposición jerárquica por partes en las explicaciones funcionales. Davies elabora sobre estos puntos y hace una teoría que llama explícitamente de capacidades sistémicas, que no pretende ser una alternativa a la propuesta de Cummins sino una versión más detallada. Más adelante expondré más a detalle los elementos de una teoría de capacidades sistémicas. La idea intuitiva, es que para entender cómo un sistema lleva a cabo una actividad, una explicación funcional divide al sistema en subsistemas y analiza cómo cada componente contribuye causalmente para que se realice dicha actividad. Como se verá, un elemento importante, sólo sugerido por Cummins, pero defendido explícitamente por Davies es que estas funciones no son postuladas arbitrariamente por el investigador, sino que su postulación obedece a la naturaleza del fenómeno; para referirme a esta característica utilizo el término *identificadas*. El resultado es parecido a la idea de las teorías teleológicas de función de que las funciones son "descubiertas."

La tesis resultante, una teoría del conocimiento basada en el análisis funcional entendido como capacidades sistémicas (CCS) es muy parecida a CFP. En CCS, el conocimiento también es una creencia verdadera que fue producida por un sistema que tiene como función producir creencias verdaderas en dicho contexto. La diferencia es la teoría de función que subyace. A pesar de que considero que una teoría de capacidades sistémicas no sólo es la mejor alternativa para evaluar funciones cognitivas, sino para evaluar funciones biológicas en general, no me comprometo con esta tesis. Mi trabajo es compatible con un dualismo sobre funciones, donde algunas sean teleológicas y otras sistémicas. El compromiso es con la tesis de que las evaluaciones de funciones cognitivas sólo pueden llevarse a cabo con la segunda noción de función descrita.

4.2. Descripción de la tesis

Como primer objetivo, el alcance de CCS es la posibilidad de considerar conocimiento a creencias verdaderas sobre los objetos macroscópicos que encuentra un individuo en su rango ecológico o en ambientes cognitivos normales⁷⁵ en el caso de seres humanos (más adelante ahondaré sobre esta noción). No obstante, no hay razones para que la teoría no pueda aplicarse a otros tipos de conocimiento como aquellos derivados del testimonio o inferencias. El naturalista sobre el conocimiento debe preocuparse por generar una base suficiente de creencias que constituyan conocimiento, pero no debe preocuparse si el conocimiento no agota todo el ámbito epistémico.

Es necesario tener claro que, como señalé en el capítulo 3, el conocimiento no es el único estado que puede dar cuenta del valor epistémico. Recapitulando sobre el mismo ejemplo del capítulo 3, no es lo mismo saber que un teorema es verdadero por consultar una fuente fiable que por *comprender* el contenido del teorema, ni el grado de justificación es el mismo. Un ejemplo claro de un bien epistémico adicional al conocimiento es la certeza. El conocimiento es una creencia verdadera que no se debe a la suerte epistémica y que rastrea de manera aproximada la verdad en los escenarios que importan. Nos interesa adquirir conocimiento, es decir es valioso, debido a que basta con esto para tener una representación del mundo suficientemente adecuada para la acción y para generar nuevas creencias relacionadas con la proposición que conocemos. Del mismo modo, cuando atribuimos conocimiento nos interesa que un agente esté suficientemente facultado para estas cosas. Visto así, el conocimiento es un mínimo⁷⁶, no un máximo, que se acercaría más a lo que buscan otros estados epistémicos como la certeza. La tesis es que las funciones cognitivas

⁷⁵ Aquí uso el significado ordinario de "normal", cuando se trata del término técnico de Millikan "Normal" siempre lo utilizo mayúscula.

⁷⁶ Con ello no niego que sea el mínimo absoluto, dado que puede haber otros estados epistémicos más austeros, como la acreditación perceptual (Cfr. Burge, 2003).

bastan para satisfacer este mínimo; el objetivo aquí es usar las evaluaciones funcionales para hacer una teoría del conocimiento, no para cubrir toda la epistemología.

El conocimiento como análisis funcional de capacidades sistémicas consiste en la siguiente tesis, reformulando CFP y sustituyendo "función propia" por "función identificada":

Una creencia verdadera B es conocimiento si y sólo si fue producida por un conjunto de sistemas cognitivos de generación de creencias $S_1, S_2 \dots S_n$, que tienen la función identificada Φ —posiblemente entre otras— de producir creencias verdadera de tipo T al que pertenece B en el contexto C.

El proceso para realizar una adscripción funcional es el siguiente:

- a) Un individuo genera una creencia verdadera y adecuada⁷⁷ B en un contexto C.
- b) Se identifican los sistemas $S_1 \dots S_n$ que produjeron B.
- c) Se identifica cuál es la función de los sistemas $S_1 \dots S_n$ que tienen un papel destacado en la formación de B de acuerdo con un dominio de evaluación (implícito al hablar del tipo de creencia al que pertenece B).
- d) Si *una* de las funciones de los sistemas $S_1 \dots S_n$ es generar creencias verdaderas en C (no exclusivamente), y están ejecutando esa función en la generación de B, B es un ejemplar de conocimiento.

Como se verá, la diferencia más notoria con respecto al capítulo anterior, radica en el punto c. La identificación de una función no se hace a partir de la historia evolutiva, sino a través de un

⁷⁷ Un poco más adelante explicaré qué entiendo por una creencia adecuada.

análisis acerca de los sistemas que llevan a cabo la producción de la creencia. El punto c parece a primera vista sospechoso. ¿Cómo se identifican dichos "sistemas"? Si todo subsistema se define a su vez por su función (dentro de un sistema), suena a que puede haber un problema de regresión y/o circularidad. Más adelante al hablar sobre sistemas cognitivos aclararé estos problemas. Para hacerlo es necesario tener más claridad con respecto a la identificación de las funciones. Mi idea es presentar una teoría específica sobre cómo identificamos funciones cuando hablamos del dominio epistémico. Primero, presentaré de manera individual los elementos que se requieren.

a) **Dominio de evaluación:** Las evaluaciones epistémicas pueden entenderse como un caso especial de evaluaciones de ejecución (i.e. conocimiento como logro, Cfr. Greco, 2010), lo cual implica la existencia de un dominio de evaluación. La tesis del conocimiento como logro, como se verá a continuación, no implica que el sujeto de atribución del logro sea un agente. El dominio está constituido por un objetivo (creencias verdaderas), por las maneras en que se deben obtener esas metas (i.e. elementos contextuales de evaluación) y por una unidad de evaluación (que usualmente es un agente, aunque no en este caso).

b) **Unidad de evaluación y condiciones de satisfacción:** La unidad de evaluación en las ejecuciones epistémicas son ejemplares de ejecución de funciones cognitivas (la noción de función en c) en la adquisición de creencias. Una ejecución cumple con los criterios para ser conocimiento, si produce una creencia verdadera por un sistema que realiza su función.

c) **Identificación de funciones:** Las funciones cognitivas relevantes al conocimiento se asignan a *tipos* de sistemas generadores de creencias se llevan a cabo a partir de su función (ver d).⁷⁸

⁷⁸ Los tipos no necesariamente se limitan a una especie. Por ejemplo, para una descripción más abstracta es posible caracterizar un sistema de manera que capture las generalidades de sistemas homólogos de distintas especies. Por otra parte, se pueden caracterizar tipos de sistemas de conjuntos de individuos atípicos o incluso casos singulares.

d) **Individuación de sistemas:** Los tipos sistemas y sus respectivas funciones se individualizan y se identifican a partir de criterios sistémicos. A estas funciones asignadas de forma no arbitraria, las llamaré funciones identificadas (FI). Los criterios sistémicos tienen que ver con la organización

e) **Otros elementos contextuales de evaluación:** Los sistemas relevantes son aquellos cuya función identificada sea producir creencias verdaderas que constituyan una representación adecuada. Lo que hace a una creencia verdadera una representación adecuada en un conjunto de contextos, parte del modo como cumple una serie de metas que objetivamente corresponden a un organismo.

En los siguientes apartados daré una explicación de estos elementos y la manera como conjuntamente permiten asignar funciones cognitivas que puedan ser consideradas como conocimiento.

4.3 Epistemología de la virtud

Las evaluaciones epistémicas, son acerca de las creencias adquiridas o sostenidas por un agente. Cuando se distingue una creencia verdadera que no constituye conocimiento de una creencia verdadera que además es conocimiento, se está llevando a cabo una evaluación epistémica. Esto tiene implicaciones tanto normativas como valorativas. Adquirir o mantener el conocimiento de que P, es algo que *ceteris paribus*, es deseable para un agente epistémico. Por otra parte, saber que P es más valioso que meramente creer verdaderamente que P. Existen otros términos epistémicos evaluativos que han sido tratados ampliamente en la literatura, en esta investigación el objetivo será únicamente el conocimiento.

En la perspectiva que defiendo en este capítulo, el conocimiento será considerado como un conjunto de propiedades relacionales que posee una creencia verdadera. Esto también es el caso para las teorías expuestas en los capítulos 2 y 3. Esto implica un rechazo explícito al primitivismo, la tesis de que el conocimiento no es analizable de manera no- circular a partir de otros.

En lo que sigue, el carácter normativo de las evaluaciones epistémicas —una de las desiderata del conocimiento— se explica a partir de la tesis de que el conocimiento es un caso particular de evaluaciones de ejecución. Por lo anterior, lo que se requiere para entender el conocimiento qué constituye el dominio de evaluación, cuál es la unidad de evaluación, el contexto de metas que lo constituyen, etc.

A manera de preámbulo, hablaré un poco sobre los antecedentes de la tesis que defiendo tal como han sido desarrollados en la epistemología de la virtud fiabilista, que forma parte de la literatura de la epistemología tradicional. En el capítulo 2 rechacé el requisito de fiabilidad, entendido del modo más usual, como tasa de creencias verdaderas sobre creencias totales. Una de las razones para rechazar este requisito fue el problema de los falsos positivos. Los autores a los que haré referencia a continuación no apelan inequívocamente a este principio en los trabajos citados⁷⁹, aún si fuera el caso, no forma parte de la parte central de su propuesta. Por lo anterior, hablaré de sus trabajos quitando la fiabilidad entendida como tasa de creencias verdaderas (o si se quiere, entendiendo la fiabilidad de modo más amplio como disposiciones estables que llevan a la verdad).

Concretamente, me concentraré en la tesis del conocimiento como logro. En términos generales, una evaluación de ejecución se plantea a partir de un dominio, constituido por determinadas metas.

⁷⁹ Por ejemplo, Greco señala "Dado que las habilidades son disposiciones fiables, esta explicación del conocimiento es una versión del fiabilismo" (2010, p. 10). Pero la noción que parece estar en juego en el trabajo es la de disposiciones *estables*. Sobre la ambigüedad en los trabajos de Sosa acerca de si sostiene un requisito de fiabilidad como tasa de creencias, ver Eraña, 2009.

La evaluación determina si se consiguió una determinada meta bajo cierto contexto. La segunda parte de la propuesta, consiste en asimilar las evaluaciones epistémicas a este tipo de evaluaciones. Para ello, basta con señalar que lo epistémico constituye un dominio semejante al de los ejemplos cotidianos como dar al blanco en la arquería. Estas analogías son claras debido a que son dominios de evaluación cuyas metas son claras, no obstante queda por aclarar cómo se constituye el dominio de evaluación del conocimiento.

En epistemología esta tesis ha sido defendida por Sosa (2009), Greco (2010) entre otros autores. Esta postura ha sido denominada epistemología de la virtud fiabilista, en contraste con una epistemología de la virtud responsabilista (donde no basta que el agente actúe de manera virtuosa, sino que tiene que actuar asumiendo ciertos deberes que usualmente se caracterizan bajo la forma de reglas). Aunque se le debe a Sosa el trabajo seminal acerca de dicha tesis, sus textos se han concentrado en discutir los problemas tradicionales de epistemología (contraejemplos Gettier, escepticismo, etc), y no en explorar de manera más detallada los aspectos metafísicos y semánticos de la tesis. Sosa describe su modelo como la teoría AAA del conocimiento, una creencia es conocimiento si y sólo si es: Acertada [Accurate], Atinada⁸⁰ [Adroit], y Apta. La primera condición hace referencia a que se cumpla la meta epistémica (i.e. la verdad), la segunda a que la creencia se genere de manera hábil o virtuosa, mientras que la tercera condición de aptitud significa por definición que la creencia sea acertada porque fue atinada. Es claro que el peso teórico está sobre la segunda parte. ¿Qué hace que una creencia sea adquirida virtuosamente? Como se verá a continuación, esta pregunta tiene consecuencias importantes tanto para la posibilidad de naturalizar el conocimiento como para el resultado de las evaluaciones epistémicas.

⁸⁰ Una traducción más fiel sería "hábil." Traduzco como "atinada" para conservar la mnemotecnica de Sosa.

Cabe señalar además, que la teoría de Sosa aplica para lo que llama "conocimiento animal", donde se refiere a un tipo de conocimiento no-reflexivo y donde la evidencia no juega un papel. Un segundo *nivel* de conocimiento, consiste en que el agente tome el conocimiento animal como insumo y reflexione acerca de su situación epistémica con respecto a éste, aceptándolo conscientemente. Este dualismo tiene como resultado un fracaso parcial (o para los optimistas, un éxito parcial) si lo que se busca es defender CN. Por otra parte, la idea de Sosa de que el conocimiento reflexivo utiliza como insumos el conocimiento animal de manera iterada es incompatible con buena parte de la literatura empírica acerca del conocimiento reflexivo humano, tomando como modelo la teoría dual de procesos (Eraña, 2009). De hecho, para el naturalista que desee buscar una contraparte empírica a la tesis de Sosa⁸¹, la teoría dual de procesos ofrece posibilidades interesantes para examinar la relación entre las creencias reflexivas y las creencias no-reflexivas, aunque las conclusiones epistémico-normativas y de arquitectura de razones que se obtengan dependerán en buena medida de la propuesta empírica que uno tome (Cfr. García-Campos, 2009).

Por su parte, Greco (2010) ha desarrollado la propuesta de Sosa de un modo que hace explícitos algunos presupuestos y de forma que se evita el dualismo. Greco ha propuesto una semántica para evaluar el éxito de un agente —entendido en sentido austero, bajo la forma que más adelante llamaré "cuasi-agencia"— enfatizando que hay ciertas cosas que se espera que haga un "agente" en ciertas circunstancias. Es decir, Greco reintroduce un elemento de responsabilidad en la semántica de evaluación, aunque la responsabilidad no figura directamente al momento de hacer las evaluaciones (por ello, no importa si el agente está asumiendo dicha responsabilidad o actuando reflexivamente al momento de adquirir/aceptar creencias). Asimismo, Greco considera que si bien

⁸¹ Sosa no suele tratar el problema metodológico en sus textos, pero en una intervención durante una presentación de Kornblith de "Epistemic Agency" (IIFs-UNAM, 14 de enero 2011) señaló que su propia visión de la agencia epistémica es compatible con el naturalismo. Aún si no fuera el caso, un naturalista ontológico podría inspirarse en la epistemología de Sosa para defender una meta-epistemología.

no es necesario juzgar a todo el agente, sí debe haber un grado alto de integración cognitiva. Considero que esto último tiene consecuencias indeseables para una perspectiva naturalista. Muchas creencias de origen modular son producto de sistemas que funcionan de manera adecuada para producir dicho tipo de creencias, pero que no trabajan de forma cognitivamente integrada; es decir, sólo comparten insumos una vez terminado el procesamiento de la tarea, pero no se apoyan durante la realización de la misma.

¿Debemos entonces concederle un estatus epistémico menor que el conocimiento, tal como la acreditación? En principio hay que rechazar esta alternativa, los conocimientos que son productos de un sistema que resulta ser modular cumplen en principio con todos los desiderata del conocimiento. El objetivo que buscaré será elaborar una manera de evaluar las funciones cognitivas de los sistemas generadores de creencias en el mismo espíritu de la epistemología de la virtud fiabilista, pero evitando tanto el dualismo de Sosa como la normatividad tan fuerte impuesta por Greco.

4.4 Dominio de evaluación

A continuación hablaré en general de cómo debe llevarse a cabo una evaluación de ejecución, para cualquier dominio, cuando sea pertinente haré algunas precisiones acerca de las evaluaciones de ejecución epistémicas. Un término evaluativo primario es aquél que sólo hace referencia al logro de una de las metas del dominio y a las condiciones normales que están presupuestas. Por ejemplo, en computación "eficiente" es un término evaluativo primario, pero también podemos, por ejemplo, evaluar un programa diciendo que "ahorra costos de energía" (por ejemplo, respecto de otro programa que termina en un mayor tiempo y utiliza más recursos del sistema donde corre). En este caso, se trataría de un término evaluativo derivado.

Definiré de manera un tanto más concreta a estos términos. Propongo entender a los términos evaluativos primarios del siguiente modo. Diremos que z es un término evaluativo primario si y sólo si:

- a) El dominio x busca el objetivo y (posiblemente entre otros).
- b) La característica z se define con conceptos que *necesariamente* pertenecen al dominio x .
- c) La característica z está relacionada con lograr y .

Por otra parte, algo es un término evaluativo derivado si y sólo si:

- a) El dominio y busca el objetivo y (posiblemente entre otros).
- b) La característica z se puede definir con conceptos que no pertenecen al dominio x .
- c) La característica z está relacionada con lograr y al menos en un contexto C , y no para todos los contextos.

Un término evaluativo derivado puede ser útil al coincidir parcialmente con un término evaluativo primario, como el caso de las medidas operacionales (el CI no coincide completamente con las habilidades intelectuales que busca medir, pero las pruebas de CI no son infalibles debido a que la coincidencia no es perfecta). Por otra parte, un término evaluativo puede coincidir completamente con su término primario correspondiente relativo a un contexto. A continuación muestro un ejemplo de cómo un término evaluativo primario de la industria automotriz puede coincidir con un término evaluativo derivado, siempre que se mantengan constantes las leyes de la física (contexto).

En la industria automotriz, un término evaluativo primario es la "estabilidad de marcha." Un auto

tiene estabilidad de marcha si puede mantener su dirección adecuadamente mientras está en movimiento. Existe otro término que se suele utilizar; en su lugar la estabilidad de fuerza-ángulo, un tipo de estabilidad dinámica definida con precisión en términos de gravedad, fuerzas inerciales y arrastre aerodinámico (Peters y Iagnemma, 2009). La estabilidad de fuerza-ángulo es un parámetro que hace referencia a principios físicos, es decir, al modo como se realiza la estabilidad de marcha con las leyes físicas que existen. La estabilidad de fuerza ángulo puede medirse con precisión, utilizando el conocimiento de física que tienen los ingenieros automotrices.

Por supuesto, podría haber otras maneras de medir la estabilidad de marcha. No sé lo suficiente sobre la industria automotriz, pero imagino que podrían pedirle a un conductor experto que manejara por una pista curva esquivando conos varias veces, y sacar un promedio de éxito de sus ejecuciones. Sin embargo la estabilidad de fuerza- ángulo en cierta manera puede ser más útil. Por ejemplo, si queremos mejorar la ejecución de un modelo de auto, podemos calcular qué efecto tendrán en la estabilidad de fuerza-ángulo ciertas modificaciones en su estructura.

A partir de esta distinción, la tesis de que el conocimiento es natural (CN) puede distinguirse de una teoría del conocimiento que únicamente sea empíricamente informada, como sucede con literatura más reciente en epistemología de la virtud (Fairweather, 2014). En el caso en el que el conocimiento es una clase natural, los sistemas cognitivos aparecerán en las descripciones del conocimiento como término evaluativo primario; mientras que una teoría que únicamente sea empíricamente informada aparecerán como términos evaluativos derivados (sólo accidentalmente coextensionales).

Aquí aparece una razón decisiva para preferir CCS que CFP —y que FE, en la medida en que comparte los elementos teleológicos— cuando se trata de defender una teoría del conocimiento basada en la ejecución de los sistemas cognitivos. Supongamos por un momento que no existen

los problemas de CFP señalados en el capítulo pasado, y que la noción de conocimiento obtenida por nociones evolutivas tiene como resultado una noción que cumple con todas las características deseables para la normatividad epistémica. Aún si esto es el caso, la obtención de las metas epistémicas no es constitutiva al conocimiento-CFP. O bien el defensor de CFP acepta que el conocimiento no es un término evaluativo —una vía que debemos rechazar si queremos que el resultado de naturalizar el conocimiento conserve el trabajo teórico y práctico que realiza el término preteórico— o bien el defensor de CFP acepta que su noción de conocimiento es un término evaluativo derivado. Si esto es el caso, hay entonces un término evaluativo primario, definido a partir de las metas que constituyen al dominio epistémico (concretamente, a la parte del dominio que corresponde al conocimiento).

Quiero resaltar que esto es un punto semántico independiente del contenido de CFP. El defensor de CFP (e.g. Graham, 2014) puede hacer notar cómo su noción de conocimiento tiende a coincidir con las metas epistémicas (al menos en los escenarios normales), incluso llevando a cabo un argumento a priori (e.g. sobre cómo la selección natural trabaja explorando un espacio de diseño y llega a soluciones que cumplen con la meta epistémica), pero dado que la meta no es constitutiva del conocimiento-CFP, sigue sin haber una naturalización del término evaluativo primario. La teoría que propongo de CCS, incluye como elemento constitutivo a las metas epistémicas, pero no se limita a éstas ya que incluye elementos sobre las características que debe tener un sistema al cual le asignemos esta función (continuidad y productividad). Si esto es correcto, CCS puede tanto naturalizar al conocimiento como mantener su dimensión normativa, como término evaluativo primario.

4.5 Unidad de evaluación y condiciones de satisfacción

Lo que distingue el presente enfoque de la epistemología de la virtud, es que la unidad de evaluación no es un agente epistémico, sino a nivel de sistemas. Existen buenas razones para tomar esta decisión metodológica. El trabajo en delfines de Herman y Morrel-Samuels (1990) citado en el segundo capítulo hace referencia a casos de conocimiento donde parece que hay una noción implícita de agencia o al menos de integración cognitiva; dado que los delfines combinan información de distintos insumos para formar representaciones adecuadas acerca de un medio ambiente. Mi intención es sostener que la unidad de evaluación son sistemas, y no agentes (si bien las metas del agente están incluidas dentro de los elementos contextuales de evaluación). Por tanto, a continuación muestro las razones que existen para rechazar explícitamente una noción de agencia o un análogo. Estos son los tipos principales de agencia que tengo en mente:

- a) **Agencia reflexiva:** Esta es la noción de agencia que se utiliza en el internismo epistemológico, donde hay un agente que reflexiona acerca de qué creencias adquiere o mantiene. Los criterios que se le exigen al agente dependen de la teoría, usualmente son de coherencia o responsabilidad frente a la evidencia.
- b) **Agencia no-reflexiva:** Esto se puede ver en la noción de Burge de “agencia primitiva”, donde los componentes de un organismo trabajan para conseguir las metas biológicas de dicho organismo. (Al igual que lo que sostengo en este trabajo, Burge rechaza que dichas metas sean dadas por una historia evolutiva, para una posición contraria ver Graham, 2014). La noción de metas no es la que corresponde a un agente consciente. Del mismo modo hay una noción de agencia epistémica implícita en la teoría de *racionalidad mínima* de Cherniak (1990). En este caso la agencia está descrita de manera más abstracta, a nivel de las relaciones inferenciales en el conjunto doxástico de un agente que deben proveer la consistencia mínima para una conducta apropiada.

- c) **Cuasi-agencia:** Greco (2010), en su versión de una epistemología de la virtud, rechaza los requisitos internistas de reflexión consciente. No obstante, la unidad de evaluación que sugiere para la adquisición de una creencia son todas las partes del aparato cognitivo que de manera *integrada*, contribuyen para su generación⁸². Para lo anterior utilizo el término de cuasi-agencia, que no aparece en el trabajo de Greco.

Dedicaré un breve espacio a criticar los requisitos excesivos de la agencia reflexiva, debido a que existe literatura amplia sobre el tema; se trata de uno de los blancos que los externistas epistémicos han atacado con más éxito. Por otra parte, noción de agencia no-reflexiva, al menos en su versión de agencia primitiva, no es contraria a la dirección general de la teoría de evaluación cognitiva que propongo.

El problema con la tesis de agencia reflexiva, es que señala que para evaluar una creencia como un ejemplar de conocimiento, se está evaluando la ejecución de un individuo completo (o de su aparato cognitivo completo). Contra esto, sostengo que a pesar de que los criterios de evaluación incluyen a todo el individuo, esto no es necesario para la evaluación del funcionamiento de un sistema.

Recurro a una analogía, para ilustrar —no tanto para apoyar— la tesis que pretendo defender: la función más básica de un músculo puede ser una unidad de evaluación al mirar si cumple con ciertas normas (e.g. si se contrae con la intensidad y la dirección que debe tener, etc), sin tener en cuenta lo que hace el resto del organismo; a pesar de que dichas normas están constituidas por la estabilidad y locomoción que debe tener el organismo completo. La analogía parece un arma de

⁸² Es decir, para la generación de dicho ejemplar de creencia. Las partes del aparato cognitivo que son relevantes para la evaluación pueden variar en cada ejecución.

doble filo: alguien podría hacer notar que, aunque es posible evaluar a un músculo de manera aislada, la manera más propia de decir que un músculo funciona adecuadamente es mirando cómo trabaja en coordinación con el resto del organismo. Esto es verdad, siempre y cuando no se pierda la perspectiva de evaluar la contribución individual de cada músculo. Siguiendo con la analogía, sabemos que cuando un músculo trabaja de manera inapropiada la posición de los huesos y ligamentos puede deformarse para compensar la deficiencia. En este caso, lo que observamos es un movimiento “exitoso”, pues permite al organismo desplazarse, pero aquí *ex hypothesi*, la ejecución del músculo es deficiente.⁸³

Tenemos entonces tres puntos:

1. La evaluación de una ejecución tiene como objeto el comportamiento de aquello que es causalmente prominente⁸⁴ para que se dé la ejecución.
2. Sólo los subcomponentes que tienen una contribución causal prominente son relevantes para la evaluación.
3. El nivel de granularidad adecuado para una evaluación implica distinguir entre los subcomponentes causalmente relevantes.

De estos tres puntos se deriva que la cuasi-agencia es la noción más cercana a ser la adecuada para realizar evaluaciones funcionales de un agente cognitivo. A pesar de que el dominio de evaluación

⁸³ Alguien podría insistir en que el funcionamiento del músculo es adecuado, como una especie de verdad analítica (acerca de las evaluaciones de ejecución): toda parte que conforma un sistema que cumple con sus objetivos, funciona de manera adecuada. Además de ser una falacia de composición, esta falta de granularidad acerca de la contribución causal de cada subcomponente va completamente en contra del modo como de hecho se realizan las evaluaciones funcionales.

⁸⁴ Las causas de un suceso son demasiadas, por eso es necesario restringir a aquellas causas que son prominentes. Independientemente de qué teoría de la causalidad sea la más adecuada para explicaciones en ciencia cognitiva, me parece que la noción de destacamiento causal requiere de un aparato que permita variaciones contextuales en la atribución. Bernstein (2014) ofrece una teoría del destacamiento causal inspirada en la teoría contrafáctica de Lewis (1973).

incluye metas referidas a la totalidad del individuo, en la noción anteriormente mencionada, lo que se está evaluando es un ejemplar de ejecución de un sistema y no el agente⁸⁵. Debido a que existen casos de evaluación donde incluir al agente completo —la totalidad de su aparato cognitivo— resulta superfluo, esto muestra que la agencia no es un elemento constitutivo de las evaluaciones en la adquisición de creencias.

Respecto a la cuasi-agencia, la integración cognitiva no es necesaria para que al menos algunas creencias constituyan conocimiento. De acuerdo con Greco (2010), es necesario apelar a la integración cognitiva para evitar las objeciones contra el fiabilismo usualmente llamadas el problema de los procesos efímeros y extraños. No obstante, como hace notar Montemayor (2014), existen procesos cognitivos que no se encuentran cognitivamente integrados, como reportes de creencias falsas sobre ilusiones ópticas (Cfr. Stottinger y Perner, 2006) que no se manifiestan en la manipulación motora, donde la acción es adecuada al tamaño real y no al tamaño reportado de los objetos. La integración no es una condición necesaria para que exista conocimiento, pero aun así es problemática.

No hay una razón por la cual las evaluaciones tengan que ser modulares. Las creencias donde el sistema que lo produjo involucra distintos subsistemas —posiblemente, pero no necesariamente, con uno de control central— pueden evaluarse del mismo modo que las creencias por sistemas de tipo modular, mostrando cómo la interacción de dichos sistemas trabajando de manera integrada es adecuada. Al principio de este capítulo señalé que un primero objetivo es cubrir un tipo de creencias básicas. No obstante, una creencia básica (en este sentido) no implica que ésta haya sido producida por un sistema modular. No existe ninguna razón a priori para que esto sea el caso; por

⁸⁵ Una motivación para evaluar sobre ejemplares de ejecución: el mismo sistema puede funcionar correctamente, pero puede tener fallos de ejecución en ocasiones.

el contrario, para muchas creencias existirá colaboración de distintos módulos (aunque algunas puedan corresponder al trabajo de un solo módulo).

Una vez determinada la unidad de evaluación, es necesario decir algo con más precisión sobre cuáles son sus condiciones de satisfacción. Para ello, sugiero retomar la idea de darle a las evaluaciones de conocimiento una semántica de éxito. La semántica de éxito es una idea de Ramsey (1991) que consiste en ligar al éxito con el contenido de una creencias verdaderas, que ha sido adaptada en distintos contextos tanto para defender una teoría sobre condiciones de verdad de una creencia (Whyte, 1990), como para elaborar teorías sobre el contenido de una creencia (Mellor, 2012, cap. 5) o de una creencia verdadera (Dokic y Engel, 2003).⁸⁶ Otro antecedente a esta manera de evaluar el éxito aplicado a la epistemología puede encontrarse en Montemayor, en borradores de “Success, Minimal Agency and Epistemic Virtue “ (2012) aplicaba la semántica de éxito de Ramsey a ciertos problemas relacionados con las evaluaciones epistémicas (la versión final del texto habla de éxito en términos generales sin hacer mención a dicho análisis semántico).

Como una nota final a este apartado, cabe señalar que el fiabilismo tradicional de procesos como aquél defendido por Goldman (1979, 1986) (en la interpretación recibida) comparte el aspecto de la unidad de evaluación. En primer lugar, el sistema no necesariamente es fiable en su tasa de producción de creencias cuando realiza su función identificada. En segundo lugar, no cualquier método puede definirse como un sistema productor de creencias. De modo que la teoría aquí defendida es diferente con respecto a la manera de evaluar aquello que produce una creencias (potencialmente, conocimiento), es más restrictiva respecto a qué tipo de cosas producen conocimiento (sólo sistemas cuya función identificada sea producir conocimiento (y no asignada

⁸⁶ Sobre las intenciones originales de Ramsey, Dokic y Engel (2003) defienden la tesis exegética de que Ramsey no pretendía ofrecer una teoría sobre la verdad, puesto sobre esto sostenía un tipo de deflacionismo.

al arbitrio del evaluador epistémico), es más restrictiva sobre la ontología de estos sistemas (debe ser posible llevar a cabo un análisis sistémico de los mismos), y tiene una normatividad que no está constituida exclusivamente por la verdad como meta. Dados los elementos anteriores, se puede ver las diferencias que tiene esta propuesta con respecto al fiabilismo son más grandes que las coincidencias.

4.6 Identificación de funciones e individuación de sistemas

Un análisis funcional sigue la siguiente forma general (p. Cummins, 1972, 76):

x funciona como Φ en un sistema s (o: la función de x en s es hacer Φ) relativo a una descripción [*account*] analítica A de la capacidad de s para llevar a cabo Ψ , sólo en caso de que x sea capaz de llevar a cabo Φ en s y la descripción A da cuenta de la capacidad de s para llevar a cabo Ψ , en parte, apelando a la capacidad de x de llevar a cabo Φ en s .

A lo anterior se le puede agregar la condición de que “ A especifica el mecanismo físico en s que instancia tales capacidades sistémicas.” (Davies, 2003, p. 27, Cfr. Cummins, 1975). Aplicando esta fórmula repetidas veces a las diferentes partes x de s tendremos un análisis funcional del sistema Ψ . Cabe señalar que la función Φ no es una función propia como en el caso de las funciones seleccionadas, sino que el análisis está relativizado a ella. Para cualquier sistema que tenga una disposición Φ , *en principio* es posible aplicar la forma general de Cummins, lo cual puede generar problemas.

Los análisis funcionales que tienen un interés explicativo son proporcionales a

- i) el grado en que las capacidades que analizan son menos sofisticadas que las capacidades que se analizan

- ii) el grado en que analizar capacidades que se analizan son distintas en grado a las capacidades analizadas,
- iii) la relativa complejidad de la organización de las partes [sistémicas] y procesos que componen al sistema (Cummins, 1975, p. 78).

Una sugerencia de Millikan para mostrar las limitaciones del análisis funcional es el análisis del corazón como un sistema que produce el ruido de los latidos, lo cual también constituye un análisis complejo. ¿Por qué los investigadores no suelen escoger este tipo de funciones que suenan intuitivamente arbitrarias? Lo que lleva a los investigadores a escoger una determinada Φ es el “destacamiento funcional” [*functional salience*] (Davies, 2003). Si por destacamiento funcional entendemos los intereses subjetivos del investigador, entonces el contraejemplo de Millikan resulta un motivo de preocupación para defender CN a partir de CCS. ¿Por qué a un investigador que contempla el latido del corazón con un espíritu naturalista tendría que interesarle más la circulación de la sangre que el sonido que produce?

La solución de Davies es inclinarse por una teoría objetiva acerca del interés teórico y el destacamiento funcional. al ofrecer como constreñimiento la organización jerárquica. "Los sistemas que son estructuralmente jerárquicos son *estructuralmente cohesionados*, donde un cambio en la organización a nivel inferior afecta el ejercicio de la capacidad a nivel superior. (2005, p. 83). De acuerdo con esta visión, las explicaciones paradigmáticas sobre funciones buscan explicar este tipo de sistemas, conectando la contribución causal de sus componentes estructurales con su aportación causal a un sistema de orden superior del que forman parte (otros constreñimientos posibles han sido provistos por teorías de función más afines a la teoría general de sistemas, donde la autopreservación de los ciclos lleva a cabo el constreñimiento). Confirmar esta tesis requeriría reunir numerosos estudios de caso sobre asignaciones de funciones en general, y concretamente

en el caso de la ciencia cognitiva, por lo que sólo la ofrezco aquí como una respuesta plausible al problema de la proliferación.

Además, quisiera explorar otro constreñimiento que puede ser añadido: el de un interés teórico unificador. Tomemos la tesis de que en biología es posible identificar *tipos de rasgos* (alas) o *tipos de funciones* (volar), en casos donde éstas no son homólogas (i.e. no desciende de un rasgo ancestro que aún llevara a cabo la función). Si la historia evolutiva no permite unificar dichos rasgos, ¿qué nos autoriza a hacerlo? Mathen (2009) sugiere que en el caso de las alas, una disciplina independiente (la aeronáutica), es la que nos permite unificar a las alas y a su respectiva función como tipo. Es decir, un rasgo análogo debe ser "definido por una ciencia independiente de modo que es reconocible en realizaciones homólogas" (p. 362) e igualmente aplicable a realizaciones no-homólogas. En el mismo trabajo, Mathen critica la idea de que la racionalidad se pueda naturalizar, dado que considera a la lógica como el mejor candidato para unificar las distintas realizaciones; sin embargo la evidencia empírica muestra que el razonamiento —incluyendo el razonamiento humano— no converge hacia la lógica.⁸⁷ Independientemente de que se pueda satisfacer el constreñimiento de Mathen para una noción vaga como la racionalidad⁸⁸, considero que se puede satisfacer para el conocimiento, aunque de una manera oblicua.

Desde la perspectiva del proyecto naturalista de CN, el estudio empírico del conocimiento no cuenta con una ciencia que lleve a cabo este papel auxiliar —la epistemología tradicional es sólo un estudio de nuestra noción preteórica de conocimiento y no una ciencia independiente. No obstante, hay un interés teórico que sirve para comenzar a unificar dichas funciones por un proceso

⁸⁷ Al menos hacia la lógica clásica. ¿Podría darse el caso que exista una lógica no-monotónica que sea un estándar normativo de razonamiento desde el punto de vista epistémico y que además modele el razonamiento típico en seres humanos? Quizá, si nos limitamos a la parte deductiva. Pero presumiblemente muchas de las inferencias tanto normativa como descriptivamente adecuadas incluyen tipos semánticos que no pueden ser modelados fácilmente como reglas.

⁸⁸ Tal como se usa el término, la racionalidad es una propiedad que pueden tener las creencias en grados; mientras que el conocimiento proposicional se adscribe de manera absoluta.

de ensayo y error. El investigador busca entender cómo es que distintos organismos representan el medio ambiente de forma adecuada para lograr sus objetivos. Este interés lleva a postular sistemas y funciones del aparato cognitivo como hipótesis, que requieren de confirmación empírica (y que evidentemente son sujetas a ser rechazadas). Los trabajos citados de etología cognitiva en el capítulo 2 son un ejemplo de cómo se puede llevar a cabo este trabajo, posiblemente con un estudio complementario con mayores controles donde dichas hipótesis puedan ser refutadas (e.g. un estudio de laboratorio llevado a cabo por psicólogos comparados). Si todo sale bien, eventualmente contaremos con criterios independientes para unificar al conocimiento como un rasgo análogo. Esto puede dar una respuesta al problema de la proliferación que satisface los requisitos de Mathen, y en conjunto con otros constreñimientos (como el de organización jerárquica) puede proveer un programa de trabajo para estudiar el conocimiento como clase natural.

A continuación tomaré algunos elementos sobre la caracterización de mecanismo de Craver y Darden para mostrar cómo se pueden reconocer sistemas cognitivos con sus respectivas funciones. Puesto que el término "mecanismo" suele connotar aquellas partes del aparato cognitivo que funcionan de manera modular, en lo que sigue mantendré este uso (sin el requerimiento de que sean innatos). Mi objetivo es mostrar cómo se pueden individuar sistemas en este sentido, para posteriormente extender esto a sistemas (conjuntos de subsistemas integrados) y finalmente a procesos donde intervienen sistemas distintos. De acuerdo con Craver y Darden (2013 p. 76), un mecanismo consiste en entidades y actividades constituidas por lo siguiente:

- a) Entidades y actividades
- b) Condiciones de inicio y terminación.
- c) Continuidad productiva.
- d) Regularidad

- e) Organización (espacial, temporal, activa).
- f) Niveles de sistemas.
- g) Contexto de operación.

Puede parecer que en principio las explicaciones como las de la etología cognitiva no son propiamente mecánicas, debido a que no siempre existe una descripción mecánica-neurofuncional. No obstante, en la segunda parte de este capítulo mostraré cómo al menos algunas de las explicaciones de la etología cognitiva pueden entenderse como esquemas de explicación mecánica, en otros casos como proto-esquemas y en ocasiones no son explicaciones mecánicas en absoluto sino que son descriptivas, no obstante describen conductas que deben juzgarse a partir de una evaluación funcional (y por ello, evalúan implícitamente sistemas, a pesar de que dichos sistemas no forman parte de la descripción). Un esquema de explicación mecánica se distingue de una explicación mecánica completa en la medida en que no hay una descripción completa de la contribución causal de todas las partes y subpartes para la producción de las condiciones de terminación del sistema. En algunos casos, las explicaciones etológicas no llegan a ser esquemas, sino proto-esquemas, debido a que no sólo son esquemas a los que le faltan pasos.

Una ventaja de la descripción de niveles de Craver y Darden es que dichos niveles están constituidos por la organización mecánica definida a partir de las funciones. Es decir, no se trata de una distinción de niveles ontológicos, sino de una organización jerárquica, sólo en los niveles más bajos hay un cambio en el nivel de explicación.

El contexto de operación tiene que ver con la organización temporal y jerárquica de los submecanismos que constituyen a un mecanismo para producir las condiciones de terminación de dicho sistema. En general, para fines del conocimiento, los distintos sistemas cognitivos deben entenderse en el contexto de sistemas productores de conocimiento. A su vez, los sistemas

productores de conocimiento deben entenderse en un contexto más amplio (lo cual se deriva de las condiciones de éxito). Por ejemplo, cuando un sistema produce creencias que serán utilizadas posteriormente como insumos de un sistema sensimotor, esto también es parte del contexto mecánico bajo el cual se juzga la función de dicho sistema. Es decir, al menos para fines del conocimiento, a pesar de que los sistemas relevantes son cognitivos, el contexto de operación dentro del cual su función se explica puede incluir sistemas de otro tipo.

La organización espacial no implica que los sistemas sean localizables, ni siquiera cuando corresponden con mecanismos (Bechtel y Richardson 2010), lo cual es especialmente relevante en el caso de los sistemas cognitivos, no sólo en el caso de los procesos integrados (e.g. los módulos no tienen que ser físicamente localizables, aunque frecuentemente estarían asociados a una red neuronal fija). La organización temporal a su vez no implica que los sistemas tengan que ser secuenciales. La organización puede ser por medio de ciclos donde los insumos de un mecanismo pueda actuar como productos para otro, y a su vez este otro subsistema produzca nuevos insumos para el primero.

Una vez que tenemos estos desiderata para identificar mecanismos, podemos utilizar exactamente los mismos para identificar sistemas, relativizándolos al contexto de un organismo llevando a cabo un tipo de tarea. Ilustraré la propuesta con un ejemplo, el giro fusiforme es un correlato anatómico importante del reconocimiento de objetos en general. Presumiblemente existe un *mecanismo* implementado en esta parte del cerebro cuya *función identificada* es llevar a cabo una tarea específica relacionada con el reconocimiento de objetos. Existe también evidencia de imagenología de que el giro fusiforme es un correlato anatómico de la capacidad de reconocer rostros (Gauthier et. al., 2000)⁸⁹, o de manera más detallada "el sistema nuclear consiste de tres parches que se activan en

⁸⁹ Los criterios de individuación de mecanismos y sistemas que propongo no requieren que haya correlatos neuroanatómicos claramente identificables. Esto es una particularidad del ejemplo.

ambos hemisferios cerebrales y que se localizan en las circunvoluciones occipitales inferiores (IOG), la circunvolución fusiforme lateral (LFG) y el surco temporal superior (STS)" (García-Aguilar, 2013). Ahora, supongamos que además de este mecanismo, el reconocimiento facial utiliza o bien otros mecanismos de dominio específico (i.e. sólo sirven para reconocer rostros), o bien un conjunto de mecanismos que —sean o no de dominio específico— trabajan conjuntamente sólo para identificar rostros. Una manera de describir una hipótesis —que describo sólo con fines de ejemplificar— con respecto a lo que sucede en este caso es que hay un mecanismo que lleva a cabo determinadas tareas de reconocimiento, que juega un papel como parte de un sistema de reconocimiento facial. El sistema tiene todas las características de un mecanismo, pero sólo de manera relativa a llevar a cabo dicha actividad (que consiste en un conjunto de tareas).

4.7 Ejemplos de aplicación y objeciones

Cabe notar que este modo de caracterizar al conocimiento, a diferencia del fiabilismo ecológico (FE) propuesto por Kornblith, no se compromete con que corresponda con un único término teórico correspondiente a una disciplina. La noción de conocimiento que se desprende de CCS permite un pluralismo metodológico que no está comprometido con una sola disciplina ni con sólo un nivel de descripción; estos últimos dependerán de las disciplinas que entren en juego en la explicación (Cfr. García-Aguilar, manuscrito). Al mismo tiempo, CCS provee un mapa para buscar una unidad teórica detrás de un conjunto de fenómenos cognitivos que tienen como resultado creencias. Este pluralismo se da no sólo en el hecho que se pueden caracterizar distintos sistemas como parte de distintas disciplinas, sino que un mismo sistema se puede describir apelando a diversas disciplinas, tal como sucede con la creciente integración entre la etología cognitiva y la psicología comparada. De hecho, estas dos disciplinas son un excelente ejemplo de complementariedad. La etología cognitiva, con una metodología de estudio al natural, ayuda a notar

conductas regulares producto de representaciones donde se pueden identificar funciones candidatas. A su vez, la psicología comparada con los controles de laboratorio adecuados puede realizar una reconstrucción más detallada sobre el sistema que lleva a cabo dicha función (si es el caso) y los mecanismos que la componen. Podemos añadir a la neurociencia cognitiva como una disciplina que puede detallar la organización espacial de los mecanismos que componen al sistema y que la conecta con su contexto de operación.

En el caso del conocimiento humano, la naturalización del conocimiento a partir de evaluaciones de función con capacidades sistémicas incluye, por supuesto, casos paradigmáticos de sistemas generadores de conocimiento. Carey y Spelke (Carey y Spelke, 1996) describen un número pequeño de sistemas separables de dominio específico a los que llama conocimiento nuclear [core knowledge]. Estos sistemas de "conocimiento" son cuatro o posiblemente cinco (Spelke, 2000): representación de objetos, representación de agentes sociales, representación de números, representación del medio ambiente y —discutiblemente— identificación de miembros del mismo grupo social. Estos sistemas se distinguen por tener cada uno límites característicos⁹⁰, por ser universales⁹¹, por ser innatos⁹² y además por ser similares a los que se encuentran en otras especies, particularmente en primates (presumiblemente, homólogos).

Pero CCS no se limita a los sistemas que cumplen con estos constreñimientos tan estrechos. CCS permite considerar como conocimiento a creencias producidas por sistemas que no son universales, como en el caso de los individuos neurodivergentes. Los casos expuestos en el capítulo 3 sobre reconocimiento facial en individuos con trastorno de personalidad límite y

⁹⁰ Por ejemplo, en el caso de la percepción de objetos, el número máximo de objetos que se pueden rastrear al mismo tiempo.

⁹¹ La evidencia proviene de los estudios que hace la antropología cognitiva, de grupos sociales muy distintos como los Piraha en el Amazonas. (Spelke, 2000).

⁹² Estas habilidades con los límites característicos respectivos a cada una, se encuentran presentes en infantes.

razonamiento espacial en individuos en la parte funcional del espectro autista, en la medida en que cumplen con las características para individuar un sistema (continuidad productiva, etc) pueden ser sujetos de evaluaciones de ejecución. Puesto que CCS tampoco incluye un requisito de que los mecanismos sean innatos, es compatible con las hipótesis neuroconstructivistas.

El conocimiento-CCS cuenta con los componentes necesarios para llevar a cabo las funciones evaluativas que se esperan del conocimiento, pero no es una descripción puramente abstracta como el fiabilismo de virtud al tener un compromiso ontológico con la existencia de sistemas organizados con las funciones identificadas apropiadas. Estos requisitos ontológicos hacen que CCS sea una tesis más sustanciosa que lo que sería simplemente una epistemología de la virtud empíricamente informada.

Hasta este momento, me he concentrado en mostrar cómo CCS resuelve los problemas de CFP y cómo resulta compatible con el estudio empírico del conocimiento. A continuación expondré algunas objeciones que considero que son particularmente problemáticas para CCS:

a) Contraejemplos Gettier: Es posible construir contraejemplos Gettier del conocimiento-CCS. De manera análoga al problema señalado por Plantinga (1993) a su propia teoría del garante epistémico a partir de funciones propias⁹³, yo podría tener la creencia de que tengo un auto Chevrolet a partir del correcto funcionamiento de mi aparato cognitivo según los criterios de CCS. Sin que yo lo sepa, el auto fue robado y desvencijado, pero por coincidencia y también sin que yo lo sepa, me gané un Chevrolet en la rifa de la oficina. Por tanto, tengo una creencia verdadera que fue producida por un sistema realizando su función identificada, pero se debe a la suerte epistémica.

⁹³ Recordemos que en el caso inusual de Plantinga, la función propia no es provista de manera selectiva sino apelando al diseño inteligente.

b) Imposibilidad de atribución de conocimiento: Así como CFP hace el requisito implausible de que los sujetos que hacen atribuciones de conocimiento tengan que conocer la historia evolutiva del rasgo involucrado (ver capítulo 3), CCS hace un requisito análogo al pedir que conozcan los detalles del funcionamiento del sistema que produce la creencia en cuestión. (La objeción está basada en un argumento de Braddon-Mitchell y Jackson (2006) en contra de las teorías teleológicas del contenido mental).

c) Individualismo: Los sistemas considerados por CCS pertenecen a individuos. La teoría tiene un compromiso con el individualismo epistémico, pero existen buenas razones para pensar que hay una dimensión social fuerte en el conocimiento (e.g. agencia epistémica colectiva). Por lo tanto, CCS no puede dar cuenta de la dimensión social del conocimiento. Asimismo, CCS parece excluir medios de transmisión del conocimiento, tal como el testimonio.

d) Imposibilidad de mejoramiento epistémico: Una de las tareas de la epistemología es ayudar al mejoramiento epistémico. Pero CCS solamente evalúa si los sistemas generadores de creencias cumplen con su función identificada, sin iluminar acerca de cómo pueden ser otras maneras mejores de adquirir conocimiento.

e) Artefactos: Los seres humanos habitualmente generamos creencias ayudados por artefactos, mismos que no forman parte de nuestro aparato cognitivo. Por tanto CCS excluye una gran parte de las creencias verdaderas que cumplen con los desiderata del conocimiento.

f) Compensación: Puede haber casos donde un sistema cognitivo puede dar productos deficientes debido al malfuncionamiento, pero el mecanismo que recibe esos productos como insumos puede llevar a cabo una compensación, como sucede en el síndrome de Williams

(Karmiloff-Smith et al., 1997). De este modo, el sistema puede dar como resultado creencias verdaderas obtenidas de manera virtuosa, pero sin que haya un funcionamiento adecuado. Por tanto CCS no es el caso.

g) Fragmentación: Puede haber situaciones donde dos o más sistemas, trabajando de forma modular, compiten por controlar el comportamiento del organismo (Carruthers, 2006). De manera análoga, los sistemas relevantes pueden competir por generar una creencia (como representación estable), de una forma en que el sistema con la respuesta más "convinciente" prevalezca. En este caso hay un proceso que no es fortuito, pero no se cumple con el requisito de cuasi-agencia impuesto por CCS.

h) Artificialidad: Se pueden realizar evaluaciones cognitivas de actividades epistémicas que se alejan del funcionamiento natural del aparato cognitivo, como cuando se estudia la cognición de los jugadores de ajedrez. Dado que se pueden evaluar funciones cognitivas en contextos artificiales, CCS no captura una noción natural de conocimiento⁹⁴.

En la lista anterior no he considerado objeciones provenientes del internismo epistemológico, dado que considero que son problemas comunes a otras teorías externistas. A continuación mostraré brevemente qué líneas generales puede seguir la respuesta a cada una de estas objeciones. Se trata sólo de un esbozo para mostrar que es posible contestar a la objeción de un modo plausible. Las respuestas detalladas tendrán que ser motivo de trabajos posteriores.

⁹⁴ De suyo no es un argumento contra la teoría, pero constituye un problema para las motivaciones que me llevaron a proponer CCS en el contexto más amplio de este trabajo.

Para responder al problema contraejemplo Gettier, Plantinga propone que el fracaso se debe a que el aparato cognitivo del sujeto funciona correctamente en el ambiente⁹⁵ *global* (en un ambiente normal donde los objetos grandes suelen permanecer en donde están), pero no en un ambiente *local* (el barrio donde los autos son robados y desvencijados). Dado que la propuesta no tiene criterios claros acerca de cuándo es válido relativizar el medio ambiente a las condiciones locales, esta línea de respuesta corre con el riesgo de enfrentarse a un problema análogo al de la generalidad para el fiabilismo clásico. Mi respuesta es que la teoría defendida por CCS no es un análisis por condiciones necesarias y suficientes, sino que es una propuesta para unificar un conjunto de fenómenos cognitivos como "conocimiento" en condiciones normales, que son aquellas en las que los mecanismos operan de manera regular y continua. Las situaciones como aquellas descritas por los casos Gettier no están cubiertas por la teoría, y puesto que la metodología que defiende no está basada en el análisis conceptual de una noción preteórica del conocimiento, no son propiamente un contraejemplo. Aun así, hay razones para pensar que CCS en ocasiones podría encontrarse en línea con las nociones preteóricas de conocimiento a este respecto. Que nuestra inclinación sea no atribuir conocimiento, no significa que esto sea producto de una evaluación epistémica (donde el sujeto no cumple con ciertos estándares), así como cuando decimos que un cerdo no tiene la capacidad de volar no está reprobando ningún tipo de evaluación. Esto es consistente con el hecho de que en los casos Gettier no hay un fallo por parte de los sujetos o de los sistemas que forman parte de ellos.

La objeción de Jackson a la teleosemántica que transporto a las atribuciones de conocimiento y que mencioné brevemente en el capítulo 3, puede responderse con los recursos que da el mismo Jackson para mostrar que las teorías de contenido de rol causal no presentan el mismo problema. De acuerdo con esta línea de respuesta, las atribuciones de contenido ordinarias o folk capturan

⁹⁵ Es parte de lo que he llamado contexto.

de manera burda el aspecto de rol causal (por ejemplo, al atribuirle una creencia P a un sujeto, esperamos que se comporte en concordancia con P y que tenga otras creencias que son consecuencia de P, etc). El mismo problema llevado a las atribuciones de conocimiento no tiene que ver con el contenido; presuponemos que de algún modo podemos determinar el contenido de la creencia, pero el problema es determinar si ésta es conocimiento. La respuesta consiste en que nuestras atribuciones de conocimiento ordinarias incluyen de manera burda las nociones de buen funcionamiento. Una manera de reunir evidencia en favor de esto es por medio de filosofía experimental, pero aquí lo dejaré como una hipótesis plausible.

Sobre la objeción de que CCS implica una epistemología individualista, cabe señalar que la noción de cuasi-agencia defendida en este capítulo no es exclusiva de individuos. Un ejemplo donde podemos ver un sistema colectivo, funcionando de manera integrada, es en los ciervos de cola blanca. En esta especie, durante el forraje, cuando un individuo se pone en estado de alerta (por escuchar un sonido, etc) levanta la cola, lo cual produce una señal, dado que se muestra el reverso, de color blanco. Si un cierto número de individuos muestran la alerta, la manada realiza una acción evasiva. A pesar de que cada ciervo de la manada tiene una creencia respectiva de que hay un peligro inminente (una vez que se tomó la decisión colectiva de huir), la creencia es producto de un sistema distribuido. Lo que se presenta de manera colectiva es la "sensibilidad epistémica" (Palermos y Pritchard, 2018).

Para responder al problema del mejoramiento epistémico, señalaré dos cosas. Por un lado, como señalé en el capítulo 3, el conocimiento es un mínimo, no un máximo. Aún si no fuera posible indicar mejoras epistémicas al conocimiento, esto no implica que no sea posible el mejoramiento epistémico de ningún tipo. A una creencia que constituye conocimiento se le pueden añadir garantías, justificación, el agente puede aumentar su comprensión acerca de la proposición que cree, etc. Por otra parte, considero que CCS sí deja lugar para plantear mejoras epistémicas al

conocimiento. El funcionamiento de los sistemas generadores de creencias puede ser virtuoso pero aun así subóptimo. El estudio de dichos sistemas puede mostrar hacia qué dirección apunta un funcionamiento óptimo. En ocasiones las mejoras se podrán implementar fácilmente, cuando intervengan sistemas que sean cognitivamente penetrables (sistemas de reflexión y monitoreo) o cuando las capacidades de discriminación se puedan adquirir por medio del aprendizaje/entrenamiento. El estudio empírico del conocimiento-CCS puede mostrar qué tipo de habilidades es conveniente adquirir de forma que sirvan de la mejor manera con insumos para uno o más sistemas del aparato cognitivo. En otras instancias, los óptimos no serán implementables sólo en la etapa de desarrollo. En otros casos, no serán implementables en seres humanos —salvo en escenarios futuristas.

En el caso de los artefactos, cuando exista suficiente integración cognitiva (incluyendo sensimotora, en los casos en que sea relevante) se puede incluir al artefacto como parte de uno o más sistemas. Esto resulta menos extraño si la hipótesis de que el aparato cognitivo humano naturalmente es capaz de este tipo de acoplamientos (Cfr. Clark, 2010). Por otra parte, no hay nada que obligue a que todos los conocimientos sean de primer orden. Existen conocimientos inferenciales y conocimientos en los que interviene el aprendizaje, de modo que podemos describir de la manera usual los escenarios donde un instrumento —que no está cognitivamente integrado— juegue un papel relevante. Por ejemplo para un físico usando un contador de Geiger, podemos describir la situación epistémica en términos tradicionales (el físico cuenta con una base de creencias previas acerca de la función del aparato, a partir de las cuales puede generar creencias sobre la radiación de su objeto de estudio junto con la lectura del aparato).

La objeción acerca de la compensación de funciones parte de una falacia de composición. El hecho de que una de las partes de un sistema, tomada de manera aislada, no ejecute correctamente su función no implica que el sistema como un todo no pueda ejecutar una función adecuadamente.

Si bien la tesis de que el funcionamiento correcto de los subsistemas se hereda hacia el sistema, la tesis inversa no tiene que ser el caso.

El problema de la fragmentación surge en los casos en que diversos módulos compitan por dar una respuesta y no exista un módulo encargado de escoger (Carruthers, 2006). No tengo una respuesta clara a esta objeción, que depende de que la hipótesis de modularidad masiva sea correcta o al menos que se den escenarios de este tipo. Una solución podría ser rechazar que en estos casos haya conocimiento y atribuir un estado epistémico menos estructurado, como la acreditación epistémica. Otra posible solución consiste en cambiar una noción de integración que requiera interacción entre las partes (como podría ser alguna versión de la teoría dual de procesos), y sostener de que se trata de un sistema donde la decisión de qué módulo proveerá la respuesta final se encuentra distribuida. El problema radica en cómo unificar ese sistema, pues si apelamos a nociones de diseño habrá que admitir que es necesario aceptar funciones teleológicas al menos para algunos sistemas. Por otra parte, si apelamos al éxito de dichos módulos cuando operan de manera simultánea, la teoría se acercará cada vez más a un fiabilismo de procesos.

Finalmente, la objeción a partir de la posibilidad de realizar evaluaciones de conocimiento en escenarios artificiales, descansa sobre un error categorial. Que una actividad o dominio reúna sistemas que en un nicho natural no trabajen juntos, no implica que la evaluación sea sobre un proceso artificial. El aspecto naturalista de la teoría tiene que ver con que se evalúe el funcionamiento de los sistemas en los contextos donde trabajan de manera regular, productiva e integrada. A diferencia de CFP, en ningún CCS momento requiere que el sujeto de atribución sea un "buen salvaje" epistémico. La posibilidad de realizar evaluaciones epistémicas en todo tipo de escenarios es una virtud de la teoría, no un defecto.

4.8. Conclusión del capítulo

El objetivo de este capítulo ha sido motivar una teoría del conocimiento que pretende retomar la tesis del conocimiento como función implícita en el fiabilismo ecológico del capítulo 2 y de las teorías teleológicas del capítulo 3; pero dejando de lado tanto los elementos fiabilistas como el compromiso con una teoría seleccionista para asignar funciones propias. Mi intención ha sido proponer donde se haga justicia a la tesis de la naturalidad del conocimiento y en la cual se mantengan las funciones evaluativas de dicho término.

En este capítulo propuse que si tomamos a las funciones cognitivas como capacidades sistémicas, es posible tener los elementos necesarios para identificar funciones, adscribir las y evaluarlas, lo cual es un mínimo indispensable. Pero además, cuando se trata de capacidades que explican el éxito en la representación adecuada, dichas capacidades permiten establecer fundamentos naturales para el conocimiento.

He mostrado cómo la relación entre la evaluación de funciones cognitivas y las evaluaciones epistémicas es mucho más estrecha de lo que podría parecer: muchas de las funciones que identificamos poseen una semántica de evaluación que contiene elementos epistémico-normativos. Esto, por supuesto, es una cuestión de hecho. No se puede determinar a priori que los seres humanos y otras especies sean sujetos de conocimiento, es una cuestión de hecho, como la posesión de cualquier capacidad que no sea trivial. La evidencia empírica de distintas tradiciones y programas metodológicos —aunque no siempre integrados— muestra que, al menos para una gran cantidad de casos, la conducta exitosa de los individuos es producida por capacidades estables jerárquicamente organizadas. A su vez, la intuición original del naturalismo ontológico de que hay una estrecha relación entre el conocimiento y el medio ambiente en el que operan las capacidades cognitivas queda vindicada, y además, esta relación es constitutiva: el ambiente es el contexto de

operación que parcialmente determina las funciones de los sistemas cognitivos. Esto es un resultado más fuerte que lo que sucede en el fiabilismo ecológico y en las teorías teleológicas, donde la relación de las funciones cognitivas con el ambiente es indirecta, a través de la evolución biológica.

Las objeciones que presento en el último apartado muestran, por un lado, la necesidad de elaborar con más detalle la teoría. Por otra parte, me parece que en los esbozos de respuesta que realicé se alcanzan a ver buenas posibilidades de refinar CCS para dar cuenta de estos problemas y de otros que no haya previsto. No obstante, el lector es quien debe tener la última palabra sobre qué tan prometedor es el enfoque que propongo.

5. CONCLUSIÓN

Desde un punto de vista abstracto, la esta investigación ha consistido en una crítica a dos propuestas de naturalización (el fiabilismo ecológico y el conocimiento como función propia), y la defensa de una tercera (el conocimiento como función de capacidades sistémicas). Mi trabajo se enfoca en mostrar los puntos de desacuerdo, haciendo énfasis en las diferencias y en los problemas respectivos que ocasionan a cada postura. Ahora es momento de hacer una recapitulación acerca de su cercanía, más allá del hecho obvio de que defienden la misma tesis general.

En primer lugar, las tres propuestas están elaboradas alrededor de una familia de nociones de manera implícita o explícita: sistemas, funciones y ambientes. En segundo lugar, las tres familias de teorías están motivadas por el mismo espíritu: que hay un fenómeno que podemos identificar bajo el término "conocimiento", que nos permite unificar lo que tienen común los modos completamente diversos como una especie cognitivamente sofisticada representa su medio ambiente para navegarlo exitosamente, y que el conocimiento humano no es distinto en tipo a este fenómeno, sino un ejemplar más del mismo.

Por otra parte, la discusión que desarrollada aquí acerca de la naturalidad del conocimiento es paralela a mucha de la literatura en epistemología tradicional de corte externista, donde se parte de que una teoría del conocimiento debe explicar en qué consiste una relación adecuada con el modo como se producen las creencias verdaderas y las condiciones en que son producidas (con grandes desacuerdos acerca del papel que lleva a cabo el agente epistémico). Hay pues, un conjunto de problemas, intuiciones y propuestas comunes a quienes trabajamos en una tradición naturalista y a quienes llevan a cabo estas reflexiones en el plano puramente conceptual.

Por lo anterior, considero que el panorama intelectual sobre el conocimiento se encuentra suficientemente maduro para que en un futuro no muy lejano se pueda llegar a un consenso —en la medida en que se pueda hablar de consenso en filosofía— acerca de la naturaleza (y naturalidad) del conocimiento. Para ello, este trabajo busca mostrar que la pregunta por su estatuto metafísico —la pregunta sobre la naturalidad— no es una reflexión secundaria, sino que es urgente continuar abordándolo por su papel central para aclarar las cuestiones propiamente epistemológicas. Visto en este contexto, el objetivo de este trabajo no es sólo defender una postura particular acerca de la metafísica del conocimiento, sino realizar aportaciones críticas a esta discusión.

A lo largo de este trabajo he querido explorar las alternativas más plausibles para naturalizar el conocimiento en un sentido relevante, haciendo ver cuáles deberían ser las características que debe contar el resultado final de una naturalización exitosa. El conocimiento, entendido como algo natural, por un lado debe consistir en algo más que una explicación sobre cómo éste se realiza de manera física; por otra parte, debe mantener las características normativas y ser capaz de llevar el trabajo evaluativo que se espera de dicho término (lo que motivó en primer instancia adoptar una postura no-eliminativista).

En el segundo capítulo he mostrado que el fiabilismo ecológico no cumple con lo mínimo deseable, concentrándome en la versión que presenta Hilary Kornblith. Aunque no comparto el compromiso que tiene con el argumento que llamé metodológico (el cual indica que el conocimiento naturalizado debe identificarse con el término teórico de la etología cognitiva), esto no resultó ser el problema fundamental de su teoría. La razón principal que me llevó a rechazar su propuesta fue la parte fiabilista. A pesar de que Kornblith considera que su postura resuelve el problema de la generalidad, la historia selectiva no es capaz de resolver la subdeterminación de una manera no-arbitraria. En esta misma línea, se puede ver que la fiabilidad —entendida como

tasa de creencias verdaderas sobre falsas— no parece tener un lugar en las funciones cognitivas tal como las estudian diversas disciplinas.

El fiabilismo ecológico también tiene problemas relacionados con su apelación a la historia evolutiva como elemento constitutivo del conocimiento, pero dado que comparte estos problemas con las teorías de función propias que tratado en tercer capítulo. Ahí muestro algunas virtudes de un enfoque libre de los constreñimientos metodológicos de Kornblith, lo cual incluye la posibilidad de utilizar un pluralismo metodológico para entender mejor el conocimiento como objeto de estudio. Es decir, no es necesario constreñirse a la etología cognitiva ni a ninguna disciplina en particular. Mostré de manera muy breve cómo este pluralismo se puede ver en el caso del comportamiento antipredatorio, y en última instancia busco conservarlo para la teoría que propongo en el último capítulo. No obstante, las teorías teleológicas de función —basadas en la selección natural— en última instancia son insostenibles, aún libres de los requerimientos de fiabilidad.

Las líneas de razonamiento sobre cómo la aptitud genética no produce necesariamente sistemas que cumplan con las propiedades epistémico-normativas es común en la literatura. Busqué ir más allá del lugar común utilizando argumentos más específicos. Adoptando una perspectiva del gen como unidad de selección y centrándome en las teorías etiológicas de función, mostré cómo el conocimiento como función propia no es suficiente ni necesario para cumplir con los desiderata establecidos en el primer capítulo, mismos que aceptan prácticamente todos los naturalistas. En el primer caso utilicé argumentos familiares para mostrar la zanja entre aptitud y adaptación. Aquí mi presupuesto de que el gen es la unidad de selección fundamental juega un papel importante. El segundo caso presento argumentos que son menos comunes en el contexto de la epistemología, haciendo referencia a cómo el desarrollo y la maladaptación pueden producir sistemas generadores de conocimiento bajo estándares razonables. El defensor del conocimiento como función propia

puede mantenerse firme en su postura y sostener que estos casos no son auténtico conocimiento, pero tendrán que decir algo más para descartar los casos que menciono para que su rechazo no sea una simple petición de principio. Mis argumentos no dependen estrictamente de que el gen sea la unidad de selección, y pueden sofisticarse más para incluir consideraciones epigenéticas y de otro tipo, pero anticipo que puede haber respuestas interesantes desde una teoría de funciones propias desde la biología evolutiva del desarrollo (evo-devo).⁹⁶

La teoría del conocimiento como función identificada de capacidades sistémicas que propongo en el cuarto capítulo evidentemente no surge en el vacío, sino que busca intentar salvar el papel constitutivo de las evaluaciones de funciones cognitivas en el conocimiento. Del fiabilismo ecológico rescato la idea de que el conocimiento debe tener un componente normativo de manera constitutiva, aunque rechazo que dicho componente sea el fiabilismo, al menos entendido como un veritismo *simpliciter*. Si bien algunos pasajes de Kornblith sugieren que tiene una normatividad más fuerte en mente no la hace explícita, y en cambio en otras ocasiones explícitamente declara su cercanía con el fiabilismo de corte tradicional. Una de las cosas que busco en el cuarto capítulo es darle contenido a una normatividad que vaya más allá de una tasa de creencia verdaderas, y una vez que cuento con un esquema general sobre las evaluaciones cognitivas, muestro cómo es fácil aplicarlo a las funciones identificadas. Aprovecho para mencionar que mi visión sobre las evaluaciones cognitivas no fue elaborada *ad hoc* para la teoría del conocimiento como capacidad sistémica, sino que considero que es independiente incluso a la distinción entre naturalismo y no-naturalismo. Dada la importancia de establecer qué es lo que constituye a las evaluaciones cognitivas, dediqué un apartado a detallar los elementos más importantes. Por supuesto hace falta un trabajo mucho más detallado para construir una semántica sobre las atribuciones de conocimiento, así como incluir otros debates relevantes en epistemología (e.g. el debate entre

⁹⁶ Para una perspectiva de las funciones cognitivas desde el evo-devo, ver Griffiths (2007).

contextualismo e invariantismo), pero considero que el esquema que propongo provee de un plan de trabajo para nuevas investigaciones sobre este punto.

La diferencia más grande es que la referencia a funciones seleccionadas se abandona por completo para sustituirlas por la noción de función de la tradición sistémica. Para ello, definí a un sistema por medio de su organización funcional, donde lo funcional a su vez depende de la organización causal de contribuciones jerárquicas entre los componentes y los subcomponentes para producir algo de manera regular y continua en una clase de contextos delimitada. Estos sistemas poseen desde un punto de vista objetivo dicha función (sin importar si dicho sistema ha sido seleccionado por llevar a cabo esta función). En el caso particular del conocimiento, propuse que se trata de una propiedad que tienen las creencias verdaderas producidas por sistemas cuya función es producir creencias verdaderas y adecuadas, donde lo adecuado tiene que ver con la habilidad de los individuos para satisfacer sus metas "objetivas". Estas metas objetivas no requieren que los individuos tengan intenciones conscientes de satisfacer dichas metas, sino que se derivan de su constitución biológica (como la necesidad de evitar obstáculos, de obtener energía, etc). Aquí el lector puede sospechar que al hablar de se trata de un adaptacionismo encubierto. La discusión semántica sobre si esto es una "adaptación" o no, resulta estéril. Basta con hacer notar que las objeciones del capítulo anterior no se le aplican a esta postura.

La búsqueda de funciones epistemológicas en el sentido anteriormente descrito permite descubrir una organización sistémica en los sujetos de conocimiento, donde los componentes y subcomponentes contribuyen causalmente para producir creencias verdaderas de manera regular en los contextos adecuados. Las objeciones al final del capítulo muestra que existen diversos elementos sobre los cuales la teoría debe trabajar para fortalecerse; particularmente en temas que no se tratan de manera específica, como las atribuciones de conocimiento. Espero también que al mostrar que las funciones identificadas pueden ser fructíferas para la naturalización del

conocimiento, esto contribuya a reconsiderar su papel en las funciones cognitivas, y a su vez a éstas en los problemas tanto de epistemología como de contenido mental con los que se ha enfrentado la filosofía en las últimas décadas.

El conocimiento tal como lo concibe CCS, es natural en la medida en que las funciones identificadas capturan las relaciones causales y la organización del sistema al cual se asigna la función. Quizá esto no cumple con los criterios más fuertes de naturalización impuestos en el capítulo 1. La flexibilidad para individuar sistemas podría dificultar considerarlos como clases naturales desde un punto de vista esencialista, y quizá incluso con las clases naturales boydeanas. Pero formulando la tesis de naturalidad en términos lewisianos, el conocimiento —como propiedad de determinadas creencias— tiene un grado de naturalidad que no resulta trivial. Esto se puede ver cuando hacemos la hipótesis de que algo es una función cuando percibimos destacamiento funcional: esta hipótesis es derrotable, pues la regularidad de operación no garantiza que existan capacidades organizadas jerárquicamente, y por ende, una función. En otras palabras, la identificación de funciones y la postulación de sistemas tiene un carácter de hipótesis científica y no es una mera construcción ad hoc.

La última pregunta que quisiera abordar es si existe el conocimiento. Por supuesto, sólo una tesis de filosofía puede terminar preguntándose si el objeto que aborda realmente existe. Es decir, cómo se puede saber si el conocimiento natural cuya estructura metafísica establecí a lo largo de estas páginas, verdaderamente corresponde a un fenómeno real. Una vez que contamos con una manera sensata acerca de cómo naturalizar el conocimiento y al mismo tiempo mantener su importe normativo, ¿cómo podemos garantizar que de hecho existan estas funciones cognitivas? Esto no se puede determinar a priori. Sólo el trabajo empírico puede corroborar que de hecho existan capacidades sistémicas cuyas funciones identificadas correspondan al dominio de evaluación del conocimiento, es decir, la producción de creencias verdaderas y adecuadas.

No obstante, la respuesta que sugiere el trabajo empírico disponible en distintas disciplinas y programas de investigación —algunos de ellos aquí citados— es que esto es el caso. Incluso podemos notar que estos trabajos no han sido llevados a cabo con la intención de "naturalizar" el conocimiento del que hablan los filósofos, sino que quienes han investigado las capacidades representacionales de distintos organismos han identificado sistemas cuya función es satisfacer las metas objetivas de especies e incluso de individuos cuya cognición es distinta a la típica en sus respectivas especies. Por ello, cabe esperar que distintas disciplinas continúen encontrando nuevas maneras en que las creencias en arquitecturas cognitivas diversas puedan tener la propiedad de ser conocimiento. Esto a su vez, proveerá nuevas ocasiones de profundizar desde el punto de vista epistemológico acerca de las propiedades evaluativas y normativas del conocimiento, dado que estas características son compatibles con su carácter natural.

6. BIBLIOGRAFÍA

- Ahlstrom, Kristoffer. 2008. "Epistemology and empirical investigation." *Grazer Philosophische Studien* 76-1: 109-134.
- Allen, Colin y Bekoff, Marc. 1997. *Species of Mind: The Philosophy and Biology of Cognitive Ethology*. MIT Press.
- Andrews, Kirsten. 2015. *The Animal Mind*. Routledge.
- Alston, William P. 1995. "How to think about reliability." *Philosophical Topics* 23-1: 1-29.
- Antony, Louise. 2003. "Who's afraid of disjunctive properties?" *Philosophical Issues* 13-1: 1-21.
- Aristóteles. 1995. *Física*. Gredos.
- Armstrong, David M. 1973. *Belief, Truth, and Knowledge*. Cambridge University Press.
- Audi, Robert. 1993. "Contemporary foundationalism." *The Theory of Knowledge: Classic and Contemporary Readings*, ed. B. Wadsworth, 206-213. Ceneage Learning.
- Ayala, Francisco J. 1970. "Teleological explanations in evolutionary biology." *Philosophy of science* 37-1: 1-15.
- Bekoff, Marc; Allen, Colin; Burghardt, Gordon M (eds). 2002. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. MIT Press.
- Braddon-Mitchell, David. 2008. "Naturalistic analysis and the a priori." *Conceptual Analysis and Philosophical Naturalism*, eds. D. Braddon-Mitchell, R. Nola, 23-44. MIT Press.
- Braddon-Mitchell, David; Jackson, Frank. 1997. "The teleological theory of content" *Australasian Journal of Philosophy*, 75:4: 474-489.
- Bechtel, William; Richardson, Robert C. 2010. *Discovering Complexity*. MIT Press.
- Bergmann, Michael. 2006. *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford University Press.
- Bermudez, José Luis. 2006. "Knowledge, Naturalism, and Cognitive Ethology: Kornblith's Knowledge and its Place in Nature." *Philosophical studies* 127-2: 299-316.
- _____. 2009. "Mindreading in the animal kingdom." *The philosophy of animal minds*, ed. R Lurz, 145-164. Cambridge University Press.
- _____. 2014. *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge University Press.
- Bernstein, Sara. 2014. "Omissions as possibilities." *Philosophical Studies*. 167-1: 1-23.

- Bonjour, Laurence. 2006. "Kornblith on knowledge and epistemology." *Philosophical studies*. 127-2: 317-335.
- _____. 2010. "The myth of knowledge." *Philosophical Perspectives* 24-1: 57-83.
- Boyce, Kenneth; Plantinga, Alvin. 2012. "Proper Functionalism" *The Continuum Companion to Epistemology*, ed. A. Cullison. Continuum International Publishing.
- Boyd, Richard. 1999. "Homeostasis, species, and higher taxa." *Species: New Interdisciplinary Essays*, ed. R. A. Wilson, 141-85. MIT Press.
- Brandon, Robert N; et. al. 1994." Sober on Brandon on screening-off and the levels of selection." *Philosophy of Science*, 61-3: 475-486.
- Brown, Jessica. 2012. "Words, Concepts and Epistemology." *Knowledge ascriptions*, eds. J. Brown, M. Gerken, 31-54. *Oxford University Press*.
- Buller, David. 1999. "Etiological Theories of Function: A Geographical Survey." *Function, selection, and design*, ed. D. Buller, 281-306. State University of New York Press.
- Burghardt, Gordon M. "Cognitive ethology and critical anthropomorphism: A snake with two heads and hog-nose snakes that play dead." *Cognitive ethology: The minds of other animals*, eds. C. Ristau, P. Marler, 53-90. Psychology Press.
- Campbell, Donald T. 1974. "Evolutionary Epistemology," *The philosophy of Karl R. Popper*, ed. P. A. Schilpp, 412-463. Open Court.
- _____. 1997. "From evolutionary epistemology via selection theory to a sociology of scientific validity." *Evolution and cognition*, 3-1,2: 5-38.
- Carey, Susan; Spelke, Elizabeth. 1996. "Science and core knowledge." *Philosophy of Science*, 63- 4: 515-533.
- Carruthers, Peter T. 2006. *The Architecture of the Mind*. Oxford University Press.
- Cockram, Nathan Robert. 2014. "Challenging Exclusionary Naturalism." *Studia Philosophica Estonica*. 7- 1: 1-34.
- Colyvan, Mark. 2010. "Naturalising Normativity". *Conceptual Analysis and Philosophical Naturalism*, eds. D. Braddon-Mitchell, R. Nola, 303-314. MIT Press.
- Conee, Earl, and Richard Feldman. 1998. "The Generality Problem for Reliabilism." *Philosophical Studies*, 89: 1-29.
- Cherniak, Christopher. 1990. *Minimal Rationality*. Bradford.
- Clark, Andy. 2010. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press.

- Cosmides, Leda; Tooby, John. 1992. "Cognitive adaptations for social exchange." *The adapted mind: Evolutionary psychology and the generation of culture*, eds. J. Barkow, L. Cosmides, J. Tooby, 163-228. Oxford University Press.
- Craig, Edward. 1999. *Knowledge and the State of Nature*. Oxford University Press.
- Craver, Carl F; Darden, Lindley. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. University of Chicago Press.
- Cummins, Robert. 1975. "Functional Analysis." *Journal of Philosophy* 72: 741-765.
- _____. 1991. *Meaning and Mental Representation*. MIT.
- _____. 2000. "How does it work?" versus "what are the laws?: Two conceptions of psychological explanation." *Explanation and cognition*, eds. F. C. Keil, R. A. Wilson, 117-144. Bradford.
- Darwin, Charles. 2008. *On the Origin of Species*. Oxford University Press.
- Davidson, Donald. 1982. "Rational animals." *Dialectica*, 36-4: 317-327.
- Davies, Paul Sheldon. 2003. *Norms of Nature: Naturalism and the Nature of Functions*. MIT Press.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford University Press.
- Dennett, Daniel C. *Brainstorms*. 1978. MIT Press.
- _____. *The intentional stance*. MIT press, 1989.
- _____. 1996. *Darwin's Dangerous Idea*. Simon & Schuster
- Descartes, René. 2011. *Meditaciones Metafísicas*. Alianza.
- Dokic, Jérôme; Engel, Pascal. 2003. *Frank Ramsey: Truth and Success*. Routledge.
- American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Publishing.
- Ellis, Brian. 2007. *Scientific Essentialism*. Cambridge.
- Eraña, Ángeles. 2003. *Normatividad epistémica y diversidad cognoscitiva*. Tesis doctoral. UNAM.
- _____. 2009. "El conocimiento animal y el conocimiento reflexivo. ¿ Niveles o tipos de conocimiento?." *Teorema: Revista Internacional de Filosofía*, 28-1: 100-111.
- Esfeld, Michael; Sachse, Christian. 2004. *Conservative Reductionism*. Routledge.
- Fairweather, A., Flanagan, O (eds). *Naturalizing Epistemic Virtue*. AAVV. Cambridge University Press.
- _____. 1974. "Special sciences (or: the disunity of science as a working hypothesis)." *Synthese* 28-2: 97-115.

_____. 1975. *The Language of Thought*. Harvard University Press.

_____. *Representations: Philosophical essays on the foundations of cognitive science*. Brighton: Harvester Press, 1981.

_____. 1985. "Fodor's guide to mental representation: The intelligent auntie's vade-mecum." *Mind* 94-373: 76-100.

_____. *Psychosemantics: The problem of meaning in the philosophy of mind*. The MIT Press, 1987.

García-Aguilar, Claudia Lorena. "The idea of levels of biological organization and its relationship to some mechanistic notions", manuscrito.

_____. 2013 "La homología de sistemas cognitivos: el caso del reconocimiento de caras en primates humanos y no humanos" *Ciencias cognitivas: entre la cooperación y la integración*, eds. J. González y P. King. Porrúa.

García-Campos, Jonatan. 2009. "Justificación y racionalidad desde la teoría dual del razonamiento." *Ideas y Valores*, 58-139: 61-86.

Garfinkel, Alan. 1981. *Forms of Explanation*. Yale.

Gauthier, Isabel, et al. "The fusiform 'face area' is part of a network that processes faces at the individual level." *Journal of cognitive neuroscience*. 12-3: 495-504.

Gettier, Edmund L. 1963. "Is justified true belief knowledge?." *Analysis*, 23-6: 121-123.

Glennan, Stuart. 2017. *The New Mechanical Philosophy*. Oxford University Press.

Gigerenzer, Gerd. 2008. *Rationality for Mortals*. Oxford University Press.

Godfrey-Smith, Peter. 1994. "A modern history theory of functions." *Noûs* 28-3: 344-362.

Goldman, Alvin I. 1979. "What is justified belief?." *Justification and Knowledge*, ed. G. S. Pappas, 1-23. Springer.

_____. 1986. *Epistemology and cognition*. Harvard University Press.

_____. 1993. "Epistemic folkways and scientific epistemology." *Philosophical Issues* (1993): 271-285.

_____. 1994. "Naturalistic epistemology and reliabilism." *Midwest studies in philosophy* 19, no. 1 (1994): 301-320.

_____. 2005. "Kornblith's naturalistic epistemology." *Philosophy and Phenomenological Research*, 71-2: 403-410.

_____. 2007. "Philosophical intuitions: their target, their source, and their epistemic status." *Grazer Philosophische Studien*. 74: 1-26.

Goodman, Nelson. 1954. *Fact, Fiction, and Forecast*. Harvard University Press.

- Gould, Stephen Jay, and Richard C. Lewontin. 1979. "The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme." *Proceedings of the Royal Society of London B: Biological Sciences* 205-1161: 581-598.
- Graham, Peter J. 2014. "The Function of Perception." *Virtue Epistemology Naturalized*, ed. A. Fairweather, 13-31. Springer.
- Greco, John. 2010. *Achieving Knowledge*. Cambridge University Press.
- Griffin, Donald R. *The question of animal awareness: Evolutionary continuity of mental experience*. Rockefeller Univ Press, 1976.
- _____. 2001. *Animal Minds: Beyond Cognition to Consciousness*. University of Chicago Press.
- Griffiths, Paul E. 1992. "Adaptive explanation and the concept of a vestige." *Trees of life: Essays in Philosophy of Biology*, ed. P. Griffiths, 111-131. Springer.
- _____. *What emotions really are: The problem of psychological categories*. University of Chicago Press, 2008.
- _____. "Evo-Devo meets the mind: toward a developmental evolutionary psychology." *Integrating evolution and development: From theory to practice* (2007): 195.
- Haugeland, John. 1997. *Mind Design II*. Bradford.
- Hazlett, Allan. 2010. "The myth of factive verbs." *Philosophy and Phenomenological Research* 80-3: 497-522.
- Hegel, Georg W. F. 2013. *Ciencia de la lógica*. Las Cuarenta.
- Heinrich, Bernd. 1999. *Mind of the raven: investigations and adventures with wolf-birds*. Cliff Street Books.
- Hempel, Carl G. 1965. *Aspects of scientific explanation*. Free Press
- Herman, Louis; Morrel-Samuels, Palmer. "Knowledge Acquisition and Asymmetry Between Language Comprehension and Production: Dolphins and Apes as General Models for Animals." *Interpretation and Explanation in the Study of Animal Behavior*, eds. M. Bekoff, D. Jamieson), 283-312. Westview.
- Herrmann, Esther; Call, Josep; Hernandez-Lloreda, Maria Victoria; Hare, Brian; Tomasello, Michael. 2007. "Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis." *Science*, 317-5843:: 1360-1366.
- Hobson, Art. 2013. "There are no particles, there are only fields." *American Journal of Physics* 81-3: 211-223.
- Hume, David. 2008. *Dialogues Concerning Natural Religion*. Oxford University Press.
- Jackson, Frank. 2000. *From Metaphysics to Ethics*. Clarendon.

- _____. 2006. "The Epistemological Objection to Opaque Teleological Theories of Content." *Teleosemantics*, eds. G. Macdonald, D. Papineau, 85-99. Oxford University Press.
- Kant, Immanuel. 2012. *Crítica del Discernimiento*. Alianza.
- Karmiloff-Smith, Annette. 1994. "Beyond modularity: A developmental perspective on cognitive science." *European journal of disorders of communication*. 29-1: 95-105.
- Karmiloff-Smith, Anette; et al. 1997. "Language and Williams syndrome: How intact is 'intact'?" *Child Development* 68-.2: 246-262.
- Kim, Jaegwon. 1988. "What is" naturalized epistemology?" *Philosophical Perspectives* 2: 381-405.
- _____. 1992. "Multiple realization and the metaphysics of reduction." *Philosophy and Phenomenological Research*. 52-1: 1-26.
- _____. 2005. "Laws, causation, and explanation in the special sciences." *History and Philosophy of the Life Sciences*, 27-3,4: 325-338.
- Kitcher, Philip. 1992. "The naturalists return." *The Philosophical Review* 101-1: 53-114.
- _____, 1981. "Explanatory Unification." *Philosophy of Science*, 48-4: 507-531.
- Kornblith, Hilary. 1980. "Referring to artifacts." *The Philosophical Review*, 89-1: 109-114.
- _____. 1999a. "In defense of a naturalized epistemology." *The Blackwell guide to epistemology*, eds. J. Greco, E. Sosa, 158-170. Blackwell.
- _____. 1999b. "Knowledge in humans and other animals." *Noûs*, 33-13: 327-346.
- _____. 2002. *Knowledge and its Place in Nature*. Oxford.
- _____. 2005. "Replies to Alvin Goldman, Martin Kusch and William Talbott." *Philosophy and Phenomenological Research* 71-2: 427-441.
- _____. 2006. "Reply to Bermudez and Bonjour." *Philosophical Studies* 127-2: 337-349.
- _____. 2007. "The Metaphysical Status of Knowledge," *Philosophical Issues*, 17-11: 145-164.
- Kripke, Saul A. 1980. *Naming and Necessity*. Harvard University Press.
- Kusch, Martin. 2005. "Beliefs, Kinds and Rules: A Comment on Kornblith's Knowledge and Its Place in Nature." *Philosophy and Phenomenological Research*, 71-2: 411-419.
- _____. 2013. "Naturalized Epistemology and the Genealogy of Knowledge." *Contemporary Perspectives on Early Modern Philosophy*, eds. M. Lenz, A. Waldow, 87-97. Springer.
- LaPorte, Joseph. 2004. *Natural Kinds and Conceptual Change*. Cambridge University Press.
- Lepock, Christopher. 2014. "Metacognition and intellectual virtue." *Virtue Epistemology Naturalized*, ed. A. Fairweather, 33-48. Springer.

- Lewis, David K. 1973. "Causation." *Journal of Philosophy*, 70-17: 556-567.
- _____. 1983. "New work for a theory of universals." *Australasian Journal of Philosophy*, 61-4: 343-377.
- Lorenz, Konrad. 1970. *Evolución y modificación de la conducta*. Alianza.
- Lowe, E. J. 2002. *A Survey of Metaphysics*. Oxford University Press.
- Mathen, Mohan. 2009. "Human Rationality and the Unique Origin Constraint" *Functions: New Essays in the Philosophy of Psychology and Biology*, eds. A. Ariew, R. Cummins, M. Perlman, 341-372. Oxford University Press.
- May, Joshua; Sinnott-Armstrong, Walter; Hull, Jay G.; Zimmerman, Aaron. 2010. "Practical interests, relevant alternatives, and knowledge attributions: An empirical study." *Review of Philosophy and Psychology*, 1-2: 265-273.
- Mayr, Ernst. 1961. "Cause and effect in biology." *Science*, 134-3489: 1501-1506.
- _____. 2002. *What Evolution is*. Basic Books.
- McLaren, Ian P. 1994. "Representation development in associative systems." *Causal mechanisms of behavioural development*, eds. J. A. Hogan, J. J. Bolhuis, 377-402. Cambridge University Press.
- Mellor, D. H. 2012. *Mind, Meaning, and Reality*. Oxford University Press.
- Millikan, Ruth G. 1995. "Naturalist reflections on knowledge." *The White Queen Psychology and Other Essays for Alice*, R. Millikan, 241-264. MIT Press.
- _____. 1984. *Language, Thought, and Other Biological Categories: New foundations for realism*. MIT press.
- Montemayor, Carlos. 2014. "Success, Minimal Agency and Epistemic Virtue." *Virtue Epistemology Naturalized*, ed. A. Fairweather, 67-82. Springer.
- Nagel, Ernest. 1961. *The Structure of Science*. Hackett Publishing.
- Nagel, Jennifer. 2013. "Knowledge as a Mental State." *Oxford Studies in Epistemology, Volume. 4*, ed. T. S. Gendler, J. Hawthorne, 273-308. Oxford University Press.
- Nagel, Jennifer. "Knowledge ascriptions and the psychological consequences of changing stakes." *Australasian Journal of Philosophy* 86.2 (2008): 279-294.
- Neander, Karen. "The teleological notion of 'function'." *Australasian Journal of Philosophy* 69-4: 454-468.
- Nesse, Randolph; Williams, George. 1997. "Are mental disorders diseases." *The Maladapted Mind: Classic readings in evolutionary psychopathology*, ed. S. Baron-Cohen, 1-22. Psychology Press.
- _____. 2004. "Cliff-edged fitness functions and the persistence of schizophrenia." *Behavioral and Brain Sciences*, 27-6: 862-863.

- Nisbett, Richard E; Ross, Lee. 1980. *Human inference: Strategies and shortcomings of social judgment*. Prentice Hall.
- Noble, Denis. 2008. *The Music of Life: Biology Beyond Genes*. Oxford University Press.
- Oppenheim, Paul; Putnam, Hilary. 1958. "Unity of science as a working hypothesis." *Minnesota Studies in the Philosophy of Science*, vol. 2, eds. Feigl et al., 3-37. Minnesota University Press.
- Palermos Orestis; Pritchard, Duncan. 2013. "Extended Knowledge and Social Epistemology." *Social Epistemology Review and Reply Collective*, 2-8: 105-120.
- Pernu, Tuomas K. 2009. "Is knowledge a natural kind?." *Philosophical Studies*, 142-3: 371-386.
- Perrine, Timothy. "Against Kornblith against reflective knowledge." *Logos & Episteme*, 5-3: 2014: 351-360.
- Peters, Steven C; Iagnemma, Karl. 2009. "Stability measurement of high-speed vehicles." *Vehicle System Dynamics* 47-6: 701-720.
- Plantinga, Alvin. 1993. *Warrant and proper function*. Oxford University Press.
- Platón. 2004. *Diálogos*, Gredos.
- Pollock, John L., and Joseph Cruz. 1999. *Contemporary theories of knowledge*. Rowman & Littlefield.
- Popper, Karl R. 1972. *Objective knowledge*. Oxford University Press.
- Premack, David. 1986. *Gavagai! or the Future History of the Animal Language Controversy*. MIT Press.
- Pritchard, Duncan. 2004. "Epistemic deflationism." *The Southern Journal of Philosophy*, 42-1: 103-134.
- _____. 2005. *Epistemic Luck*. Clarendon Press.
- _____. 2010. *The Nature and Value of Knowledge*. Oxford University Press.
- _____. 2011. "What is the swamping problem." *Reasons for Belief*, ed. A. Reisner,. 244-259. Cambridge University Press.
- _____. 2014. *What is This Thing called Knowledge?* Routledge.
- Putnam, Hilary. 1975. "The meaning of 'meaning'", *Minnesota Studies in the Philosophy of Science* 7: 131-193
- _____. 1982. "Why reason can't be naturalized." *Synthese*, 52-1: 3-23.
- Quartz, Steven R.; Terrence J. Sejnowski. 1997. "The neural basis of cognitive development: A constructivist manifesto." *Behavioral and brain sciences*, 20-4: 537-556.
- Quine, W. V. O. 1969a. "Epistemology Naturalized." *Ontological relativity and other essays*. Columbia University Press.

- _____. 1969b. "Natural Kinds." *Ontological relativity and other essays*. Columbia University Press.
- _____. 2013. *Word and Object*. MIT Press.
- Ramsey Frank P. 1990. "Truth and Probability." *F. P. Ramsey Philosophical Papers*, ed. D. H. Mellor, 52-94. Cambridge University Press.
- Ristau, Carolyn A. 1991. "Aspects of the cognitive ethology of an injury-feigning bird, the piping plover." *Cognitive ethology: The minds of other animals*, eds. C. Ristau, P. Marler, 91-126. Psychology Press.
- Rivas, Jesús, and Gordon M. Burghardt. 2013. "Crotalomorphism: A metaphor for understanding anthropomorphism by omission." *The Cognitive Animal: Experimental and theoretical perspectives on animal cognition*, M. Bekoff, 9-17.
- Rorty, Richard. 1981. *Philosophy and the Mirror of Nature*. Princeton University Press.
- Rorty, Richard; Engel, Pascal. 2007. *What's the use of truth*. Columbia University Press.
- Rowlands, Mark. 2009. *El filósofo y el lobo*. Seix Barral.
- _____. 2012. *Can Animals be Moral?* Oxford University Press.
- Samuels, Richard. 2002. "Nativism in cognitive science." *Mind & Language* 17-3: 233-265.
- Schaffer, Jonathan. 2009. "On what grounds what". Metametaphysics: new essays on the foundations of ontology*, eds. D. Chalmers, D. Manley, R. Wasserman, 347-383. Oxford University Press.
- Scott, Lori N.; Levy, Kenneth N., Adams, Reginald B.; Stevenson, Michael T. 2011. "Mental state decoding abilities in young adults with borderline personality disorder traits." *Personality Disorders: Theory, Research, and Treatment* 2-2: 98.
- Shettleworth, Sara J. 2010. *Cognition, evolution, and behavior*. Oxford University Press.
- Smith,Carolynn L., and Jane Johnson. 2012. "The chicken challenge: what contemporary studies of fowl mean for science and ethics." *Between the Species*, 15-1: 75-102.
- Sober, Elliott. 1980. "Evolution, population thinking, and essentialism." *Philosophy of Science*, 47-3: 350-383.
- Sosa, Ernest. 2009. *Reflective Knowledge: Apt Belief and Reflective Knowledge, Volume II*. Oxford University Press.
- _____. 2010. *Knowing full well*. Princeton University Press.
- Spelke, Elizabeth S. 2000. "Core knowledge." *American Psychologist*, 55-11: 12-33.
- Stanovich, Keith E., West, Richard F. 1998. "Individual differences in framing and conjunction effects." *Thinking & Reasoning* 4-4: 289-317.

- Stein, Edward. 1996. *Without good reason: The rationality debate in philosophy and cognitive science*. Clarendon Press.
- Stich, Stephen P. 1979. "Do animals have beliefs?" *Australasian Journal of Philosophy*, 57-1: 15-28.
- _____. *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. MIT Press.
- Stevenson, Jennifer L.; Gernsbacher Morton Ann. 2013. "Abstract spatial reasoning as an autistic strength." *PLoS ONE* 8-3: e59329.
- Stöttinger, Elisabeth; Perner, Josef. 2006. "Dissociating size representation for action and for conscious judgment: Grasping visual illusions without apparent obstacles." *Consciousness and Cognition* 15-2: 269-284.
- Talbott, William J. 2005. "Universal Knowledge." *Philosophy and Phenomenological Research*, 71-2: 420-426.
- Tarski, Alfred. 1956. "The concept of truth in formalized languages." *Logic, semantics, metamathematics* 2: 152-278.
- Tiel, Jeffrey R. 1999. "The dogma of Kornblith's naturalism." *Synthese*, 120-3: 311-324.
- Turri, John. "Epistemic Supervenience." 2010. *A Companion to Epistemology*, eds. J. Dancy, E. Sosa, M. Steup, 340-343. Blackwell.
- Tversky, Amos; Kahneman, Daniel. 1986. "Rational choice and the framing of decisions." *Journal of business*: 51-S278.
- Unger, Peter. 1968. "An Analysis of Factual Knowledge." *The Journal of Philosophy*, 65-6: 157-170.
- Van Cleve, James. 1999. "Epistemic Supervenience Revisited." *Philosophy and Phenomenological Research*, 59-4: 1049-1055.
- Van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford University Press.
- Weinberg 2011. "Experimental Epistemology." *The Routledge Companion to Epistemology*, eds. S. Bernecker, D. Pritchard, 823-835. Routledge.
- Whyte, Jamie T. 1990. "Success semantics". *Analysis*, 50-3: 149-157.
- Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.
- Wright, Larry. 1973. "Functions." *The philosophical review*. 82-2: 139-168.