



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**PROGRAMA DE POSGRADO EN FILOSOFÍA DE LA CIENCIA
CAMPO DE CONOCIMIENTO: FILOSOFÍA DE LA CIENCIA**

**¿PUEDE EXPLICARSE EL LOGRO EPISTÉMICO DEL
ENTENDIMIENTO CIENTÍFICO EN TÉRMINOS DE UNA TEORÍA
VERITISTA DE LA EVALUACIÓN EPISTÉMICA?**

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

MAESTRÍA EN FILOSOFÍA DE LA CIENCIA

P R E S E N T A:

JOSÉ ALFONSO ANAYA RUIZ ESPARZA

TUTOR:

**DR. MIGUEL ÁNGEL FERNÁNDEZ VARGAS, INSTITUTO DE
INVESTIGACIONES FILOSÓFICAS**

MÉXICO D.F., MARZO DE 2013



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

Índice

Introducción	4
1 Cuestiones Preliminares	10
1.1 Algunas cuestiones preliminares sobre los métodos de evaluación en la ciencia	10
1.2 La tesis veritista de evaluación epistémica	15
1.3 Algunos tipos de entendimiento científicos	25
1.4 El desafío de van Fraassen	27
2 Explicación veritista de algunas modalidades de entendimiento científico	33
2.1 La explicación científica	33
2.1.1 El modelo Nomológico-Deductivo e Inductivo-estadístico	33
2.1.2 La explicación por relevancia estadística	42
2.1.3 La explicación causal	49
2.2 La unificación teórica	56
3 ¿Es posible explicar en términos veritistas el logro epistémico asociado al uso de idealizaciones y modelos en la ciencia?	77
3.1 Consideraciones preliminares sobre los modelos y las idealizaciones	79
3.2 Una propuesta insatisfactoria. La importancia del principio de que no haya premisas falsas	88
3.3 La propuesta de Catherine Elgin. La interpretación no-fáctica del entendimiento científico	97
3.4 En contra de la veracidad de las leyes científicas. La postura de Cartwright	105
3.5 Una teoría alternativa de la representación	112
3.6 ¿En qué consiste el logro epistémico asociado al uso de modelos idealizados?	126
Conclusión	138
Bibliografía	145

Introducción

Esta investigación se encuentra inscrita entre dos debates distintos: el debate sobre la naturaleza del entendimiento y el debate sobre el veritismo epistémico. La parte central de esta investigación se centrará en investigar si es posible explicar en términos de una teoría veritista de evaluación epistémica distintos tipos de entendimiento científico. Una teoría veritista sostiene, en términos generales, que la creencia verdadera tiene un papel privilegiado en toda la actividad cognitiva humana. En el capítulo 1 explicaré con más detalle en qué consiste esta teoría y presentaré la versión de veritismo que adoptaré aquí.

Es importante señalar que con “entendimiento” no me refiero a una facultad cognitiva humana, sino a un logro epistémico, es decir al logro de entender por qué p . El interés de muchos filósofos contemporáneos sobre la naturaleza del entendimiento se ha incrementado notablemente en años recientes.¹ En este debate es posible identificar claramente al menos dos posturas: la primera, enarbolada prominentemente por algunos epistemólogos, sostiene que el entendimiento es una posición epistémica distinta del conocimiento; la segunda, enarbolada típicamente por algunos filósofos de la ciencia, sostiene que el entendimiento es simplemente un tipo de conocimiento, en particular conocimiento de causas.² La postura de que el entendimiento es diferente del conocimiento ha sido defendida recientemente por Katherine Elgin, Linda Zagzebski, Jonathan Kvanvig y Duncan Pritchard. La motivación de fondo de esta postura es la intuición de que entender algo involucra algo más que meramente conocer una serie de proposiciones. Por ejemplo, uno puede saber muchas cosas particulares sobre la guerra

¹ Véase por ejemplo Kvanvig (2003), Elgin (1996, 2006, 2007 y 2009), Grimm (2006), Mizrahi (2011), Lipton (2009), Zagzebski (2001), Pritchard (2009) y Riggs (2003), entre otros.

² Pritchard (2009) y Grimm (2006) presentan esta distinción. Entre los filósofos de la ciencia que sostienen que el entendimiento es un tipo de conocimiento se encuentran Lipton (2004) y Achinstein (1983). Por ejemplo, Lipton sostiene que “el entendimiento no es un tipo de súper-conocimiento, sino simplemente más conocimiento: conocimiento de causas” (Lipton, 2004, p. 30), mientras que Achinstein sostiene que “explicar involucra una intención de volver a q entendible [...]. Tal entendimiento lo interpreto como un tipo de conocimiento. Uno entiende q sólo si uno conoce una respuesta a Q , una respuesta que uno sabe que es correcta” (Achinstein, 1983, p. 23).

del Peloponeso y, sin embargo, no tener entendimiento sobre ella porque no es capaz de apreciar ciertas relaciones que se dan en la información que conoce. Por ejemplo, un sujeto que carece de entendimiento no podría explicar qué provocó la guerra, qué estaba en juego ni cuáles son las razones que llevaron a la derrota de Atenas. En otras palabras, uno puede tener una gran cantidad de conocimiento sobre algún tema y, sin embargo, no entenderlo. Esta intuición ha llevado a algunos filósofos a interpretar al entendimiento como si fuera un logro distinto del conocimiento.

Ahora bien, la postura de que el entendimiento no es un tipo de conocimiento puede sostenerse apelando al menos a dos tipos de estrategias: a) las que sostienen que el conocimiento no es suficiente para que haya entendimiento y b) las que sostienen que el conocimiento no es necesario para que haya entendimiento. Duncan Pritchard (2009) argumenta a favor de esta postura apelando a una estrategia tipo (a). Él sostiene que uno puede saber cuál es la causa de p y, sin embargo, no entender por qué p . Un ejemplo simple ilustra esto: imaginemos que mi auto se descompone, lo llevo con un mecánico y él me explica que la causa del mal funcionamiento es que las juntas homocinéticas están completamente dañadas. Imaginemos, además, que yo soy completamente ignorante de la mecánica de un automóvil. En este caso, con base en el reporte del mecánico yo adquiero conocimiento sobre la causa del mal funcionamiento de mi auto y, sin embargo, insistiría Pritchard, no tengo entendimiento al respecto. Pritchard concluye a partir de ejemplos de este tipo que tener conocimiento causal no es suficiente para entender. Él insiste en que hace falta saber de qué manera se relaciona la causa con el efecto, en este caso, de qué manera se relaciona el daño en las juntas homocinéticas con el mal funcionamiento de mi auto (*Cfr.* Pritchard, 2009, p. 38).

Por otro lado, Kvanvig (2003), Zagzebski (2001), Elgin (2006 y 2007) e inclusive el mismo Pritchard (2009) argumentan a favor de esta postura apelando a una estrategia del tipo (b). Es decir, sostienen que tener conocimiento sobre la causa de p no es necesario para tener

entendimiento de p . Pritchard y Kvanvig argumentan a favor de esta postura apelando a que la diferencia entre entendimiento y conocimiento tiene que ver con su compatibilidad con ciertos tipos de suerte epistémica. De acuerdo con Kvanvig, el entendimiento es compatible con todo tipo de suerte epistémica. Un ejemplo que presenta el propio Kvanvig supuestamente ilustra esto: imaginemos que un sujeto en una biblioteca se acerca a un estante lleno con libros sobre la guerra del Peloponeso, toma uno al azar y lo lee. El libro que tomó es un libro fiable y el sujeto forma una gran cantidad de creencias verdaderas sobre la guerra del Peloponeso. Supongamos que las creencias que formó el sujeto lo hacen capaz de apreciar relaciones explicativas complejas en el cuerpo de información: el sujeto es capaz de responder correctamente muchas preguntas sobre la guerra del Peloponeso. Pero supongamos además que el estante del que tomó el libro fiable está repleto de libros que contienen información falsa, siendo la única excepción justamente el libro que tomó el sujeto. Kvanvig sostiene que esto sería una razón suficiente para afirmar que el sujeto no tiene conocimiento, pues una de las moralejas de los casos de Gettier es precisamente que el conocimiento es incompatible con la suerte epistémica que está involucrada aquí. Sin embargo, insiste Kvanvig, el sujeto sí tiene entendimiento de la guerra del Peloponeso porque el entendimiento, a diferencia del conocimiento, sí es compatible con la suerte epistémica (*Cfr.*, Kvanvig, 2003, pp. 196-200).

Ahora bien, Pritchard sostiene que la diferencia entre entendimiento y conocimiento es mucho más compleja, aunque también tiene que ver con la suerte epistémica, Pritchard escribe: “aunque Kvanvig y otros están en lo correcto al pensar que el entendimiento es compatible con cierto tipo de suerte epistémica que socava al conocimiento, están equivocados al pensar que es compatible con todos los tipos de suerte epistémica que socavan al conocimiento” (Pritchard, 2009, p. 37). Pritchard insiste en que hay que distinguir la suerte epistémica involucrada en los contraejemplos de Gettier y la suerte epistémica que él llama ambiental. La suerte epistémica ambiental es el tipo de suerte que aparece en el caso de Goldman sobre las fachadas falsas de

graneros.³ Por otro lado, la suerte epistémica del tipo de Gettier no es de este tipo, sino que es un tipo de suerte que interviene entre los hechos y las creencias. Un ejemplo es útil para ilustrar la distinción: imaginemos que un día salgo al cine y al regresar me encuentro con que mi casa está ardiendo. Veo a un sujeto vestido como el jefe de bomberos y le pregunto por qué se está quemando mi casa y él me responde que hubo un corto-circuito. Supongamos, además, que este sujeto no es el jefe de bomberos sino alguien que va a una fiesta de disfraces y simplemente me dice a modo de mera especulación que el incendio fue ocasionado por un corto-circuito. Imaginemos, además, que efectivamente un corto-circuito es la causa del incendio de mi casa. Por lo tanto, la creencia que formo al creer en el testimonio de este sujeto es verdadera. El tipo de suerte involucrado en la formación de esta creencia verdadera es del tipo de Gettier: hay algo que interviene entre mi creencia y los hechos, en otras palabras no hay una relación causal adecuada entre el hecho y mi creencia. Consideremos ahora una situación similar, pero en la que pregunto efectivamente al jefe de bomberos y obtengo la misma respuesta. Supongamos, además, que frente a mi casa hay decenas de personas disfrazadas como jefes de bomberos (supongamos que todos van a la misma fiesta) y casualmente le pregunto al indicado. El tipo de suerte involucrado en la formación de esta creencia verdadera es del tipo de suerte ambiental. Pritchard argumenta que el entendimiento es compatible sólo con la suerte epistémica ambiental, pero no con la suerte epistémica tipo Gettier. Así, de acuerdo con Pritchard, la diferencia entre entendimiento y conocimiento es mucho más sutil de lo que había sostenido Kvanvig. Presento esta postura a modo de contraste con la postura que sostendré. Aquí argumentaré a favor de una postura contraria a la de ellos. Sin embargo, las razones a favor de mi postura surgirán de un análisis detallado de distintos tipos de entendimiento científico. Es importante presentar este panorama general de la discusión para saber en qué parte del debate se ubica la postura que defenderé aquí.

³ Goldman presenta este ejemplo en su artículo "Discrimination and Perceptual Knowledge" (1976).

Hay, sin embargo, otra estrategia para argumentar que el conocimiento no es necesario para que haya entendimiento. De manera sobresaliente Catherine Elgin argumenta que la diferencia fundamental entre entendimiento y conocimiento consiste en que el conocimiento es un logro epistémico fáctico, mientras que el entendimiento no lo es. Decir que un logro epistémico es fáctico significa que implica la verdad de lo que se cree. Por ejemplo saber que p , implica la verdad de p . Ella insiste en que el uso extendido de modelos idealizados en la actividad científica muestra concluyentemente que el entendimiento científico no es fáctico. Esta línea argumentativa típicamente ha sido ignorada por los epistemólogos que abordan el problema de la naturaleza del entendimiento.⁴ Aquí buscaré llenar este vacío. En el capítulo 3 analizaré detalladamente la postura de Elgin y argumentaré que está equivocada. Sin embargo, esto no significa que el entendimiento científico sea, sin más, identificable directamente con algún tipo de conocimiento, pues como ya vimos, hay otras razones para rechazar esta identificación, a saber las que esgrimen Kvanvig, Pritchard y compañía. Lo único que busco mostrar es que las razones de Elgin para defender la tesis de que el entendimiento es diferente del conocimiento no son buenas razones.

Sin embargo, los filósofos de la ciencia típicamente han supuesto que hay otras fuentes de entendimiento científico, además de los modelos idealizados. Tradicionalmente, por ejemplo, se ha sostenido que las *explicaciones científicas* otorgan entendimiento del mundo. Por lo tanto, el fenómeno del entendimiento científico parece ser más amplio de lo que supone Elgin. Ella insiste en que el tipo de entendimiento que surge de los modelos idealizados no es fáctico, pero no dice nada sobre los otros tipos de entendimiento científico. En esta investigación discutiré varias formas de entendimiento científico y si bien no argumentaré que todas son fácticas, sí

⁴ Por ejemplo, tanto Pritchard (2009) como Grimm (2006) no adoptan una postura sobre la propuesta de Elgin. Cuando ilustran la estrategia que insiste en que el entendimiento es no fáctico apelan a la postura de Linda Zagzebski (2001).

argumentaré que todas son explicables en términos de una teoría veritista. Como ya mencioné, ésta es la finalidad principal de esta investigación.

La estructura de la tesis será la siguiente: en el capítulo 1 expongo la versión de la teoría veritista que adopto en esta investigación, explico algunas cuestiones generales sobre la evaluación epistémica y abordo de un modo crítico la postura de van Fraassen, según la cual la explicación científica *no es* un logro epistémico de la ciencia, sino uno pragmático. En el capítulo 2 explico, desde el marco de la teoría veritista que he elegido adoptar, el valor epistémico de distintos tipos de entendimiento científico: el tipo de entendimiento que surge de distintas concepciones de la explicación científica y el tipo de entendimiento que surge de la unificación teórica. En lo que respecta a la explicación, analizo distintos modelos de explicación científica: el nomológico-deductivo e inductivo-estadístico de Carl Hempel, el modelo de relevancia estadística de Wesley C. Salmon y el modelo causal de James Woodward. En lo que respecta al entendimiento científico que surge de la unificación teórica centro mi atención en la teoría que defienden Gerhard Schurz y Karel Lambert. Finalmente, en el capítulo 3 analizo si el tipo de entendimiento científico que surge del uso de modelos idealizados es explicable en términos veritistas y argumento en contra de la postura de Elgin que esboqué anteriormente.

1. Cuestiones Preliminares

Esta investigación parte del supuesto de que el entendimiento científico es un logro epistémico propio de la actividad científica. Antes de pasar a analizar algunas formas particulares de entendimiento científico hay que decir algo sobre la evaluación epistémica en general, para explicar claramente qué quiere decir que el entendimiento científico es un logro *epistémico*. Además, como se señaló en la introducción, en esta investigación adopto el punto de vista de una teoría veritista de la evaluación epistémica, por lo tanto también es necesario explicar en qué consiste esta teoría.

1.1. Algunas cuestiones preliminares sobre los métodos de evaluación en la ciencia

De acuerdo con Larry Laudan, “en el sentido más general del término, el éxito en una actividad siempre tiene que ver con relaciones entre medios y fines [...]. Decir que una actividad es exitosa es simplemente decir que promueve los fines de [...] aquéllos involucrados en ella (o de aquellos que la juzgan exitosa)” (Laudan, 1984, p. 87). Así, por ejemplo, un día de pesca es exitoso si se atrapan suficientes pescados para alimentar a la familia completa. Desde luego, si la pesca se lleva a cabo con fines comerciales (además de los fines de supervivencia) entonces un día de pesca no es exitoso si sólo se consigue suficiente pescado para alimentar a la familia; hace falta que haya un excedente que resulte en un beneficio económico. Algo similar sucede con cualquier actividad: para evaluar adecuadamente si es o no exitosa, hay que establecer antes cuáles son los fines o las metas en relación con los cuales se llevará a cabo la evaluación. En el ejemplo del día de pesca hay que determinar cuál es la finalidad (por ejemplo, comercial o de supervivencia) para determinar si una cantidad de pescado es suficiente o no para considerar exitoso el día.

En el caso de la actividad científica ocurre algo similar. Para determinar si es exitosa o no una actividad científica particular (por ejemplo un proyecto de investigación de algún grupo de

científicos), antes hay que determinar cuáles son los fines en relación a los cuales se hará la evaluación. Pero como quedó claro en el ejemplo de la pesca, las actividades humanas pueden estar asociadas con una amplia variedad de metas y fines que varían de acuerdo con los intereses de los sujetos involucrados. Es claro que la actividad científica tiene también una amplia variedad de metas: económicas, sociales, políticas y epistémicas, entre otras. Por ejemplo, algunas compañías farmacéuticas promueven la investigación científica para alcanzar un beneficio económico; algunos Estados la promueven para alcanzar ciertos beneficios sociales o políticos y algunos científicos la llevan a cabo para descubrir la verdad sobre un fenómeno particular y entenderlo. En esta investigación centraré mi atención en las metas propiamente *epistémicas* de la actividad científica y, en particular, me centraré en el estudio de un tipo de meta epistémica de la actividad científica: el entendimiento científico.

No asumo aquí que el entendimiento científico sea la única meta epistémica de la actividad científica ni la más fundamental. Tampoco asumo que la actividad científica sólo tenga metas epistémicas. Es claro que la actividad científica también tiene metas económicas, políticas y sociales. Simplemente sucede que en este trabajo me dedicaré a la investigación del entendimiento científico, que es un tipo particular de meta epistémica propia de la ciencia. Ahora bien, se podría argumentar a favor de la tesis de que las metas epistémicas de la ciencia son las metas más fundamentales de la ciencia, en el sentido de que cualquier otro tipo de metas (por ejemplo, económicas, sociales o políticas) sólo puede alcanzarse si primero se alcanzan las metas epistémicas. Se podría argumentar, por ejemplo, que la investigación científica sólo podrá otorgar beneficios económicos si las teorías, modelos y explicaciones que usan los científicos son, por ejemplo, verdaderas o empíricamente adecuadas. Tampoco asumo que esta tesis, o alguna parecida, sea correcta, si bien la postura que defiendo aquí deja la puerta abierta a este tipo de argumentación. Es decir, no hay ningún compromiso en la investigación que presento sobre qué tipo de metas asociadas a la actividad científicas son más fundamentales.

En este punto es pertinente preguntarse cuáles son las metas propiamente *epistémicas* de la actividad científica y qué características son las que determinan cuándo una meta científica es epistémica. Estas preguntas no se responderán aquí con toda la precisión y detalle que ameritan, pues ello requeriría hacer otra investigación sobre la naturaleza de la evaluación epistémica. Adoptaré una respuesta específica a esas preguntas: el veritismo, una teoría que sostiene (en términos generales) que la meta de la actividad cognitiva en general, desde un punto de vista epistémico, es la creencia verdadera.⁵ De acuerdo con esta postura “la función primaria de la cognición en la vida humana es adquirir creencias verdaderas en lugar de creencias falsas sobre asuntos que son de interés para nosotros” (Alston, 2005, p. 29). Si se adopta una postura veritista y se asume que la ciencia es un tipo de actividad cognitiva, entonces resulta que la creencia verdadera tendrá un papel fundamental en la definición de las metas epistémicas de la ciencia. Aquí no argumentaré a favor de la corrección del veritismo. Esto se debe a que no asumo que sea correcto, sino que lo adopto como una *hipótesis* de trabajo, que debe ponerse a prueba constantemente. De hecho, en los capítulos 3 y 4 se investigará si el valor epistémico de una amplia variedad de tipos de entendimiento científico puede explicarse según la teoría veritista. Mi respuesta a esta pregunta será positiva, pero estoy abierto a la posibilidad de que haya otros tipos de entendimiento científico que sean más problemáticos para la hipótesis veritista que adopto. En la siguiente sección explicaré con mayor detalle en qué consiste el veritismo que se adopta en esta investigación.

Pero antes de pasar a esto, quiero subrayar una distinción general sobre las metas de una actividad y las metas de los individuos que la practican. Como ha quedado claro a partir de la discusión previa, el éxito de una actividad depende de las metas que se usan para evaluar la actividad. Una vez que se han establecido cuáles son las metas y finalidades con las que se va a

⁵ Esta definición del veritismo es demasiado vaga, pero en este momento se presenta sólo con fines ilustrativos. En la siguiente sección (1.2) explicaré con mayor precisión qué tipo de veritismo adopto aquí.

llevar a cabo la evaluación, resulta relativamente sencillo llevar a cabo la evaluación. Larry Laudan y Bas van Fraassen incluso llegan a sugerir que no hace falta investigar cuáles son las motivaciones de los agentes particulares involucrados en una actividad para determinar si la actividad que llevan a cabo es o no exitosa. En particular, las metas en función de las cuales se evaluará el éxito de una actividad científica no tienen que ser las metas presentes en las mentes de los científicos particulares cuando desarrollaron sus investigaciones. A diferencia de los juicios de racionalidad, sostiene Laudan, los juicios de éxito pueden llevarse a cabo sin investigar los motivos o metas del agente (*Cfr.* Laudan, 1984, p. 87). Por su parte, van Fraassen (el pionero del empirismo constructivo) insiste en que la finalidad de la ciencia consiste en proporcionar teorías que sean adecuadas empíricamente. Sin embargo, él mismo aclara que esto no significa que los científicos tengan en la mente esta finalidad al desarrollar sus teorías. De acuerdo con él, hay que distinguir entre la finalidad de una teoría y los motivos por los que un científico la desarrolla. En sus propias palabras: “la finalidad de la ciencia desde luego no debe ser identificada con los motivos de los científicos individuales. La finalidad del ajedrez es hacer jaque mate a tu oponente; pero los motivos para jugarlo pueden ser fama, dinero o gloria. La finalidad determina qué cuenta como éxito en la empresa como tal; y esta finalidad puede ser perseguida por cualquier número de razones” (van Fraassen, 1980, p. 8). Así, de acuerdo con ellos, para determinar si una actividad científica particular es exitosa o no, basta con investigar si se alcanza la meta predeterminada. No hace falta que se investigue, además, si el individuo que la llevó a cabo tenía en mente la motivación correcta al llevar a cabo la actividad.

Me parece que, a la luz de lo que se ha dicho hasta ahora, la postura que presentan Laudan y van Fraassen es correcta. Efectivamente, en la empresa evaluativa no hace falta que se investiguen las motivaciones de los agentes. Pero tiene que quedar muy claro que esto sólo es correcto si *ya* se ha establecido antes satisfactoriamente cuáles son las metas y finalidades que se usarán para evaluar la actividad. En cambio, si se busca determinar cuáles son las metas en

relación con las cuáles se debe evaluar una actividad particular, entonces sí es pertinente investigar cuáles son las motivaciones de los agentes que llevan a cabo tal actividad. En el caso de la actividad científica resultaría simplemente insatisfactorio determinar cuáles son las metas (epistémicas, sociales, económicas o políticas) de la ciencia sin tomar en cuenta las motivaciones de los científicos particulares y, especialmente, de los científicos más sobresalientes. No tomar esto en cuenta equivaldría a ignorar la larga tradición histórica de la ciencia occidental. Con esto tampoco quiero sugerir que las motivaciones de los científicos sean el único elemento que se debe tomar en cuenta para determinar cuáles son las metas de la ciencia. Así, por ejemplo, una investigación seria y profunda que busca determinar cuáles son las metas epistémicas genuinas de la actividad científica debe tomar en cuenta tanto las motivaciones de los científicos, sus opiniones sobre cuáles son las metas epistémicas de la ciencia, las motivaciones y opiniones de los filósofos de la ciencia, evidencia histórica sobre qué metas de hecho se han perseguido en el desarrollo de la ciencia, entre muchas otras. Como ya señalé anteriormente, aquí no llevaré a cabo esa tarea. Pero esa investigación arrojará el resultado de que la actividad científica ha perseguido una multitud de metas de distinto tipo. Lo que se necesita aquí es aislar las metas que sean específicamente epistémicas. Para este fin adoptaré el marco veritista, que insiste en que la creencia verdadera juega un papel fundamental en las metas epistémicas de la ciencia. De nuevo, insisto, no asumiré que es correcto, sino que lo adoptaré como hipótesis y, como tal, lo pondré a prueba en lo que respecta a su capacidad de dar cuenta del valor epistémico de distintos tipos de entendimiento científico. A continuación explicaré con mayor detalle en qué consiste el veritismo que asumo aquí.

1.2. La tesis veritista de evaluación epistémica

A pesar de que es una cuestión controvertida cuál es la estructura correcta de las teorías científicas,⁶ aquí asumiré, por simplicidad (y por compatibilidad con la literatura que se revisará) que éstas son un conjunto de proposiciones.

De acuerdo con el marco general que he adoptado aquí, una creencia (u otra actitud dóxica)⁷ es evaluada de manera positiva, *desde el punto de vista epistémico*, cuando ésta posee alguna característica o propiedad epistémicamente valiosa. El veritismo es una teoría que da un contenido específico a la noción de punto de vista epistémico. Da una respuesta particular a la pregunta: ¿Qué es aquello que buscamos fundamentalmente en nuestras empresas cognitivas? Los teóricos veritistas sostienen que la creencia verdadera es la meta central de toda empresa cognitiva y que ésta tiene un papel central al determinar qué se quiere decir con *punto de vista epistémico*.⁸ Así, de acuerdo con el veritismo, evaluar positivamente a una creencia desde el punto de vista epistémico consiste en afirmar que esa creencia tiene una propiedad que ejemplifica una relación aceptable con la verdad, llamaré a la teoría que se compromete con esta afirmación *teoría veritista de evaluación epistémica* (TVEE). Es importante notar que esta teoría veritista se centra en *el valor* epistémico de la creencia verdadera y de otros logros epistémicos. De acuerdo con la tesis central de esta teoría, en toda explicación de por qué un logro epistémico es valioso se debe apelar de algún modo al valor de la creencia verdadera. Es posible derivar una posición como ésta de las observaciones de Alston al respecto. De acuerdo con Alston, adquirir creencias verdaderas es la meta más fundamental de la cognición desde el

⁶ Ulises Moulines y José A. Díez presentan la evolución del debate sobre la naturaleza de las teorías científicas en su libro *Fundamentos de Filosofía de la Ciencia* (1997), especialmente en los capítulos 8-10. Ahí se puede apreciar que efectivamente la cuestión de cuál es la naturaleza de las teorías científicas es un tema controvertido.

⁷ En adelante se omitirá la precisión “u otra actitud dóxica”, pero el lector debe tomar en cuenta que lo que se dice aquí sobre las creencias también se aplica en principio, a otras actitudes dóxicas, por ejemplo a las aceptaciones. Cuando sea necesario hacer una distinción sobre actitudes dóxicas, ésta se hará de un modo explícito, como sucede en la sección 3.2

⁸ Véase Goldman (2001), “The Unity of Epistemic Virtues” para una defensa de un tipo de veritismo y también Goldman (1999), *Knowledge in a Social World*, especialmente el capítulo primero, para una defensa general del veritismo de una serie de objeciones en su contra y Alston (2005) *Beyond “Justification”*. Véase también Grimm (2008) “Epistemic Goals and Epistemic Values” para una discusión detallada de algunas ideas de Goldman y de Alston.

punto de vista epistémico: “no sé cómo probar que la adquisición, retención y uso de creencias verdaderas sobre cuestiones que son de interés o importancia para nosotros es la meta más básica y central de la cognición. No conozco nada más obvio de lo que podría derivarse” (Alston, 2005, p. 30). Es decir, de acuerdo con Alston, la adquisición, retención y uso de creencias verdaderas es la meta central epistémica es una cuestión casi evidente. Más adelante, cuando Alston explica el proyecto que adoptará en su investigación, sostiene que “avanzaré pensando en términos de las propiedades valiosas de la creencia, desiderata, que son valiosas epistémicamente justo en la medida en que se relacionan de alguna manera con los valores más básicos y centrales a los que se dirigen nuestras actividades cognitivas” (Alston, 2005, p. 33). Esto quiere decir que para Alston las creencias son valiosas epistémicamente sólo en la medida en que se relacionen de algún modo con la meta más básica de la cognición, es decir con la adquisición de creencias verdaderas. Puesta de este modo, la tesis veritista se reduce a lo siguiente: todo valor epistémico debe explicarse apelando a alguna relación adecuada que éste tiene con la creencia verdadera. Es decir, para que un valor epistémico sea explicado desde el punto de vista veritista basta con que exista *alguna* relación aceptable de este valor epistémico con la creencia verdadera.

Puesto de este modo, el veritismo es una teoría sobre el valor distintivamente epistémico. La centralidad de la creencia verdadera, según esta postura, radica en que ésta es necesaria e irremplazable cuando se busca explicar por qué un logro epistémico es valioso. Esta tesis es bastante fuerte, pues está cuantificada universalmente. Según ella, *todo* valor epistémico se puede explicar apelando de un modo esencial al valor de la creencia verdadera. Dar un argumento concluyente a favor de esta tesis resulta bastante complicado, dado su carácter universal. Como ya señalé antes, adopto esta tesis como una hipótesis que debe ponerse a prueba, de modo que conseguir explicar valores epistémicos en los términos de la TVEE aporta evidencia a su favor.

Ahora bien, sigue resultando ambiguo qué papel exactamente debe jugar la creencia verdadera en estas explicaciones para que la TVEE sea correcta. Para explicar con mayor detalle cuál es el papel de la creencia verdadera en las explicaciones es necesario considerar qué tesis sobre el valor epistémico subyace al veritismo.

En las teorías veritistas tradicionales, la tesis del valor subyacente suele ser muy estricta. Para ilustrarlo consideremos nuevamente la postura veritista que defiende William Alston. Según él, el veritismo es obviamente correcto, al grado que no hace falta argumentar mucho para aceptar que la búsqueda de la creencia verdadera es la característica definitoria de la empresa cognitiva humana. Al caracterizar su postura él afirma:

Sostengo que la función primaria de la cognición en la vida humana es adquirir creencias verdaderas en lugar de falsas, acerca de cuestiones que son de interés o importancia para nosotros (Alston, 2005, p. 29).

Y más adelante añade:

La idea de que es importante para el florecimiento humano guiarse por suposiciones correctas en lugar de incorrectas sobre cómo son las cosas es tan obvia que parecería innecesario insistir en ello (Alston, 2005, p.30).

Supongamos que Alston está en lo correcto y que es obvio que la verdad tiene un papel central en toda empresa cognitiva y que su valor es central cuando se trata de explicar el valor de toda actividad cognitiva. Estos fragmentos sugieren que la postura veritista particular de Alston tiene dos características específicas: que la creencia verdadera es la *única* meta epistémica *fundamental*, y que cualquier bien epistémico distinto de la creencia verdadera es bueno porque nos lleva a alcanzar creencias verdaderas. A esta postura veritista subyace la siguiente tesis acerca del valor epistémico:

Monismo-V del valor epistémico:

La creencia verdadera es el único bien epistémico fundamental (Pritchard, 2011, p. 37).

Decimos que un bien epistémico es fundamental cuando su valor epistémico no es simplemente un valor instrumental con relación a un bien epistémico ulterior, es decir un bien epistémico es fundamental cuando resulta valioso por sí mismo. Decimos que un bien epistémico es no fundamental cuando no es un bien epistémico fundamental (*Cfr.* Pritchard, 2011, p. 34). Ahora bien, si el monismo-V es un supuesto del veritismo que defiende Alston, entonces tenemos la consecuencia de que cualquier cosa diferente de la creencia verdadera que tenga valor epistémico, tiene que ser meramente instrumentalmente valioso relativo a la creencia verdadera. Es decir, cualquier bien epistémico diferente de la creencia verdadera es valioso sólo si nos lleva a conseguir creencias verdaderas.

Llamaré *duro* al veritismo que acepta la tesis monista-V del valor. Para contrastar a este tipo de veritismo con el *tolerante*, que adopto aquí, hay que señalar que el monismo-V tiene las siguientes características: a) el valor de la creencia verdadera es central (en el sentido de ser *fundamental*) y es el *único* que tiene este estatus; por otro lado, b) el hecho de que la creencia verdadera sea un bien fundamental implica que todo bien epistémico distinto de la creencia verdadera deriva su valor epistémico *únicamente* de su relación instrumental con la creencia verdadera, es decir, su valor epistémico *sólo* puede derivar del hecho de que nos lleva a alcanzar creencias verdaderas.

El tipo de veritismo que asumiré aquí se caracteriza por adoptar una tesis del valor epistémico distinta, lo llamo *tolerante* y asume como correcta la siguiente tesis del valor epistémico:

Pluralismo de relaciones del valor epistémico:

La creencia verdadera es el único bien epistémico esencial.

Decimos que un bien epistémico es *esencial* cuando no hace falta apelar a otro bien epistémico para explicar su valor epistémico. Decimos que un bien epistémico es *no esencial* si, al explicar su valor epistémico, siempre tenemos que apelar a otro bien epistémico. Esta tesis del valor epistémico también es *monista* porque, al igual que la tesis monismo-V, asume que la creencia verdadera es el *único* bien epistémico que tiene un estatus privilegiado, en este caso es el único bien epistémico esencial. Sin embargo, difiere de la tesis monismo-V en que su estatus de esencial no implica que cualquier cosa distinta de la creencia verdadera deba relacionarse *instrumentalmente* con ésta para tener valor epistémico. De aquí que sea una tesis pluralista con respecto a las relaciones que un bien epistémico distinto de la creencia verdadera debe ejemplificar con la creencia verdadera para que lo consideremos epistémicamente valioso. En otras palabras, este tipo de veritismo permite que los bienes epistémicos se relacionen de más de una forma con la creencia verdadera.

Así, esta versión tolerante del veritismo deja abierta la posibilidad de que haya una variedad indeterminada de relaciones que los bienes epistémicos pueden ejemplificar con la creencia verdadera. Más adelante presentaré qué relaciones utilizaré en esta investigación para dar cuenta del valor epistémico de las distintas formas de entendimiento científico que abordaré aquí. Por ahora es importante señalar que dar preferencia a una versión tolerante sobre una versión dura de veritismo no es una elección caprichosa ni *ad hoc*. La versión dura del veritismo es demasiado estrecha, dado que sólo da cabida a una relación con la creencia verdadera, i.e. a relación instrumental de conductividad a la creencia verdadera. Algunos problemas se derivan de su aceptación y, en su lugar, estos problemas se ven aliviados si se adopta una versión tolerante. Para ilustrar esto con un ejemplo, a continuación mostraré,

siguiendo la argumentación de Duncan Pritchard (2011), que la adopción de un veritismo duro (y su tesis monista-V del valor asociada) lleva al llamado “problema de la absorción” del valor del conocimiento.

Supongamos que la fiabilidad (*i.e.* la tendencia de un proceso a arrojar un número mucho mayor de creencias verdaderas que de creencias falsas) es una propiedad epistémicamente valiosa. Supongamos además que nos enfrentamos a dos creencias verdaderas idénticas, con la única diferencia de que una de ellas fue formada de modo fiable mientras que la otra no. Si sólo valoramos epistémicamente a la fiabilidad porque nos lleva a formar creencias verdaderas, deberíamos de ser indiferentes sobre cuál creencia tiene mayor valor epistémico, pues ambas son verdaderas. Como en ambas creencias ya está presente el valor de la verdad, no debería importarnos cuál se formó mediante un proceso fiable y, consecuentemente, deberíamos creer que ambas creencias son igual de valiosas. Así, parece que en esta situación el valor de la fiabilidad de las creencias es *absorbido* por el valor de su verdad. La tesis general axiológica presente en el problema de la absorción es la siguiente:

Tesis axiológica general:

Si una propiedad (como la de estar formada de manera fiable [...]) sólo es instrumentalmente valiosa en relación con un bien ulterior (por ejemplo, la creencia verdadera [...]), entonces de ello se sigue que, en los casos en los que ya se encuentra presente el bien ulterior, ningún valor adicional es conferido por la presencia de la propiedad que es instrumentalmente valiosa (Pritchard, 2011, pp. 31-32).

Ahora bien, esta situación representa un problema si queremos defender la intuición pre-teórica de que el conocimiento es más valioso que la mera creencia verdadera. Supongamos, junto con una teoría fiabilista del conocimiento, que el conocimiento es la creencia verdadera formada de un modo fiable. Si aceptamos la tesis monista-V del valor tenemos que todo bien

epistémico diferente de la creencia verdadera es valioso sólo si se relaciona instrumentalmente con la creencia verdadera. De este hecho y de la tesis axiológica general, se sigue que cualquier valor epistémico de cualquier cosa distinta de la creencia verdadera es absorbido por el valor de la creencia verdadera. Así, la intuición de que el valor que tiene el conocimiento sobre la creencia verdadera está dado por la fiabilidad implícita en aquél resulta insostenible, pues este valor es absorbido por el valor de la creencia verdadera. Esto nos deja con la consecuencia anti-intuitiva de que el conocimiento no es más valioso que la creencia verdadera. Éste es el problema de la absorción del valor conocimiento.

Debe quedar claro que una condición que debe darse para que se presente el problema de la absorción es que cualquier bien epistémico distinto de la creencia verdadera sea valioso meramente por ser instrumentalmente valioso con respecto a la creencia verdadera. Si se rechaza la tesis monista-V, entonces se deja abierta la posibilidad de que el valor epistémico de bienes distintos de la creencia verdadera esté dado por algo diferente de su conductividad a la verdad. Así, el problema de la absorción sugiere que es mejor adoptar una postura veritista tolerante, pues ésta no adopta la tesis monista-V del valor que da lugar al problema de la absorción.

Así, hay razones independientes para adoptar una postura tolerante y no una dura del veritismo. Sin embargo, es claro que hay que decir algo más para que la tesis tolerante del veritismo no sea una mera estipulación. Es decir, ¿cuáles son las razones, más allá de su aparente obviedad, como parece sugerir Alston, para aceptar que la verdad es la meta epistémica *esencial* de la empresa cognitiva humana? ¿Por qué todo bien epistémico debe tener alguna propiedad que lo relacione de manera adecuada con la verdad? Ya lo dije anteriormente, pero en este punto es importante insistir en que la TVEE debe entenderse como una *hipótesis*. El apoyo a su favor proviene de su capacidad de explicar el valor de una amplia variedad de bienes epistémicos en los términos que esa hipótesis establece. En la medida que sea capaz de

hacer esto de una manera aceptable, la hipótesis será aceptable. Ni Goldman (2001, 1999) ni Alston (1993, 2005) aceptan explícitamente que el estatus de la tesis veritista de evaluación epistémica (TVEE) sea efectivamente el de una hipótesis. Sin embargo, la manera en la que argumentan a favor de la corrección del veritismo concuerda bastante con lo que sugiero aquí que debe hacerse para mostrar la fortaleza de la hipótesis: analizar casos particulares de bienes epistémicos y mostrar que es posible explicar su valor en los términos de la TVEE.⁹ En buena medida esta tarea es llevada a cabo en el trabajo de Alston. Él analiza varios bienes epistémicos y muestra que su valor es explicable en términos de la TVEE. Por ejemplo, él argumenta cómo se pueden explicar, desde un marco veritista, el valor de la conductividad a la verdad, el de las características deontológicas de la creencia y el de las características de los sistemas de creencias que están entre las metas de la cognición (por ejemplo, la explicación, el entendimiento, la coherencia y la sistematicidad).¹⁰

En esta investigación me enfocaré en dar una explicación del valor epistémico de distintos tipos de entendimiento científico. Alston señala que todas las características epistémicamente valiosas de los sistemas de creencias ejemplifican una relación de *dependencia* con la creencia verdadera. Es decir, el valor de la coherencia, la sistematicidad, el entendimiento y la explicación *depende* de que la mayoría de las creencias que componen el sistema sean verdaderas o bien tengan alguna propiedad epistémicamente valiosa explicable en términos de la TVEE, por ejemplo, que hayan sido producidas mediante un método fiable o que se hayan formado de un modo epistémicamente responsable. Aquí, sin embargo, no me limitaré al uso de esta relación de dependencia. Para explicar el valor de las distintas formas de entendimiento científico que se explorarán aquí haré uso de tres relaciones. Para explicar en

⁹ Véase Fernández (2011) para una articulación más detallada de qué está involucrado en tratar al veritismo como una hipótesis.

¹⁰ Para la explicación detallada de Alston sobre cómo el valor epistémico de todos estos bienes puede explicarse en términos de una teoría veritista véase especialmente el capítulo 3 de su libro *Beyond "Justification"* (2005).

qué consiste la primera relación hace falta introducir el término de “unidad de evaluación”. Hay una variedad de cosas que pueden evaluarse desde un punto de vista epistémico: actitudes dóxicas, métodos, prácticas, rasgos de los sujetos e incluso a los sujetos mismos. Para referirme genéricamente a alguna de estas cosas utilizaré el término “unidad de evaluación”. La primera relación que presento puede ser ejemplificada por una variedad de unidades epistémicas, por lo tanto esta noción será útil para explicar en qué consiste la relación. A continuación presento estas relaciones:

1) *Dependencia*. Decimos que una unidad de evaluación ejemplifica la relación de dependencia con la creencia verdadera cuando su valor depende de que otras creencias (u otras actitudes dóxicas) sean verdaderas o bien tengan un valor epistémico distinto de la verdad que sea explicable en términos de la TVEE.

Para ilustrar en qué consiste esta relación presentaré algunos ejemplos. Decimos, por ejemplo, que un sistema de creencias coherente es epistémicamente valioso si las creencias que lo conforman son verdaderas, son altamente probables o están justificadas. Por otro lado, decimos que la actividad de razonar utilizando argumentos sólidos es epistémicamente valiosa porque las premisas de este tipo de argumentos deben ser verdaderas.¹¹ Tanto el sistema coherente de creencias como razonar usando argumentos sólidos ejemplifican la relación de dependencia con la creencia verdadera.

2) *Ser conducente a la verdad*. Decimos que un método de formación de creencias ejemplifica la relación de *conductividad* con la creencia verdadera cuando éste es fiable, es decir cuando nos lleva a formar un mayor número de creencias verdaderas que de creencias falsas; o bien cuando nos lleva a

¹¹ Aunque es cierto que el valor epistémico de razonar sólidamente no se agota aquí. Hay al menos otra relación con la creencia verdadera, a saber la relación de ser conducente a la verdad.

formar un mayor número de creencias que tienen un valor epistémico distinto de la verdad que sea explicable en términos de la TVEE, que de creencias que no lo tienen.

Por ejemplo, típicamente la percepción es un mecanismo fiable de formación de creencias. Por otro lado, un argumento sólido arroja siempre como conclusión proposiciones verdaderas, de modo que creer las conclusiones de argumentos sólidos es también un proceso fiable de formación de creencias. Ambos procesos de formación de creencias ejemplifican la relación de ser conducentes a la verdad.

3) *Ser altamente probable dada la evidencia.* Decimos que una creencia (u otra actitud dóxica) ejemplifica esta relación cuando, dada cierta evidencia, es altamente probable que la creencia (o actitud dóxica) sea verdadera.

Es importante distinguir esta relación de la propiedad de ser altamente probable simpliciter. Un ejemplo ayudará a clarificar esta distinción. La probabilidad simpliciter de que cualquier persona sufra un paro cardíaco $P(c)$ se calcula estimando la frecuencia con la que ocurren los paros cardíacos en una muestra representativa. Supongamos que $P(c)$ es de .01. Ahora bien, es cierto que hay algunos factores que se sabe que tienen una influencia considerable en la ocurrencia de paros cardíacos, como tener cierta herencia genética, la dieta y la frecuencia con la que se hace ejercicio. Para calcular la probabilidad de que cualquier persona tenga un paro cardíaco, dada la evidencia (e) de que tiene una dieta alta en grasas saturadas, raramente hace ejercicio y tiene ascendencia con un historial de problemas cardíacos hay que estimar la frecuencia de que ocurran paros cardíacos en una muestra representativa. Supongamos que $P(c|e)$ es de .7. Es importante notar que $P(c|e)$ es distinta de $P(c)$.¹² Cuando se evalúa si una

¹² Es claro que también la probabilidad de $P(c)$ se determina tomando en cuenta cierta evidencia, sin embargo el punto relevante es que la evidencia que se utiliza para determinar esta probabilidad es distinta de la evidencia que se utiliza para determinar la probabilidad de $P(c|e)$. Una notación alternativa para

creencia ejemplifica la relación de ser altamente probable dada cierta evidencia es claro que hay que considerar cual es su probabilidad $P(x|e)$, no su probabilidad $P(x)$.

En este trabajo haré uso de estas relaciones para explicar el valor epistémico de los distintos tipos de entendimiento científico que abordaré en esta investigación. A continuación señalaré cuáles son estos tipos.

1.3. Algunos tipos de entendimiento científico

La meta de esta investigación es aportar evidencia a favor de la idea de que es posible entender el logro epistémico del entendimiento científico desde un punto de vista veritista. Para hacerlo analizaré algunos tipos prominentes de entendimiento científico y buscaré explicar su valor epistémico en términos de la TVEE. Como sólo elegiré una muestra de tipos de entendimiento científico, surge naturalmente la pregunta ¿con base en qué criterios se elegirán los tipos de entendimiento científico que analizaré aquí? Utilizaré tres criterios para seleccionar qué tipos de entendimiento científico analizaré: 1) relevancia en la tradición de la filosofía de la ciencia, 2) relevancia en el debate contemporáneo sobre el entendimiento científico y, finalmente, 3) dificultad para explicar el tipo de entendimiento científico en términos de la TVEE. Elijo estos criterios para ser lo más abarcante posible en una investigación limitada como ésta.

En el capítulo 2, particularmente en la sección 2.1 examinaré con detalle tres modelos de explicación científica: el modelo nomológico-deductivo e inductivo-estadístico; el modelo de relevancia estadística y el modelo causal. Señalaré cómo efectivamente la tradición filosófica ha insistido en la idea de que las explicaciones científicas exitosas nos dan entendimiento de los fenómenos que explican. También señalaré cómo en cada caso los defensores más relevantes de estos tipos de explicación insisten en que su concepción hace justicia a esta intuición general que relaciona a la explicación científica con el entendimiento científico. En la sección 2.2 del

referirnos a $P(c)$ y que deja claro este punto sería " $P(e|c)$ ".

mismo capítulo exploraré una propuesta que presenta una forma de entendimiento científico que hace uso de la noción de unificación teórica. Esta propuesta tiene una mayor relevancia en el debate contemporáneo sobre el entendimiento científico. Sin embargo, todas las formas de entendimiento científico que analizaré en este capítulo pueden explicarse en términos de TVEE de un modo relativamente sencillo.

En capítulo 3 exploraré una forma de entendimiento científico que surge de los así llamamos modelos científicos. A diferencia de los tipos de entendimiento científico explorados en el capítulo 2, la dificultad de explicar el valor de este tipo de entendimiento en términos de la TVEE será mucho mayor. Sin embargo, lograr explicar su valor en términos de la TVEE sería una importante contribución a favor del veritismo, precisamente debido a la dificultad que representa. Además, este tipo de entendimiento científico tiene una gran relevancia en el debate contemporáneo.

Pero antes de pasar a analizar estos tipos de entendimiento científico me parece pertinente hacer una breve digresión para echar un vistazo a la postura general de van Fraassen con respecto a cuáles son las metas epistémicas y cuáles las metas pragmáticas distintivas de la ciencia. Hago esta digresión porque van Fraassen *parece* negar algo que yo asumo y que tiene mucha importancia en el capítulo 2, a saber que la explicación científica posee un valor *epistémico*: el entendimiento científico que genera; mientras que él argumenta que la explicación científica *no* es una meta epistémica, sino pragmática. Si está en lo correcto, entonces el análisis de modelos explicativos que se lleva a cabo en el capítulo 2 descansaría en un supuesto erróneo. La meta de la digresión será mostrar que este problema es una mera apariencia, pues la explicación científica *sólo* puede tener el valor pragmático que van Fraassen identifica si ésta descansa en una teoría que sí tiene propiedades epistémicamente valiosas desde el punto de vista del veritista.

1.4 El desafío de van Fraassen

Comúnmente se traza una distinción entre los logros epistémicos y los logros pragmáticos de la actividad científica. En la primera categoría suelen incluirse, entre otras, la representación correcta de los fenómenos, el poder de unificación y el poder explicativo; en la segunda suelen incluirse el poder predictivo, la manejabilidad computacional de una teoría, entre otras. Van Fraassen, sin embargo, defiende un modo muy distinto de trazar esta distinción.

Según él, la actividad científica sólo tiene una meta *epistémica* genuina: la adecuación empírica. La meta epistémica que los realistas científicos defienden como la más relevante es, desde su punto de vista, una meta espuria. Me refiero a la representación correcta (verdadera) de los fenómenos del mundo. Según la caracterización del realismo científico que presenta van Fraassen, “la ciencia pretende darnos, en sus teorías, una historia literalmente verdadera de cómo es el mundo; y la aceptación de una teoría científica involucra la creencia de que [la teoría] es verdadera” (van Fraassen, 1980, p. 8). A diferencia de este tipo de realismo, van Fraassen defiende el empirismo constructivo. Éste sostiene que la aceptación de una teoría involucra la creencia de que ésta es adecuada empíricamente y no de que es verdadera. Van Fraassen ofrece una variedad de razones para rechazar la tesis realista. Pero quizás la razón más prominente consiste en señalar que el empirismo constructivo que él defiende es igualmente útil que el realismo y que, además, tiene la ventaja de no comprometerse con la existencia de las entidades inobservables postuladas por muchas teorías científicas. Por lo tanto, el realismo constructivo tiene la ventaja de adoptar una ontología más simple que el realismo científico. De acuerdo con el realismo constructivo de van Fraassen, la meta de la actividad científica es elaborar teorías empíricamente adecuadas y no teorías verdaderas. Para entender con mayor profundidad en qué consiste la propuesta de van Fraassen es necesario explicar qué es la adecuación empírica. El propio van Fraassen afirma que “una teoría es adecuada empíricamente exactamente si lo que dice sobre las cosas observables y eventos en el mundo, es verdad –

exactamente si “salva los fenómenos”” (van Fraassen, 1980, p. 12). De manera concisa: la adecuación empírica es la verdad sobre los aspectos observables del mundo.

Así, para van Fraassen la meta epistémica fundamental de la actividad científica es la adecuación empírica. Lo más interesante de su postura es que presenta una clasificación de las metas cognitivas y pragmáticas de las teorías científicas que difiere de la que he asumido aquí. van Fraassen insiste en que la adecuación empírica es la *única* meta propiamente epistémica de las teorías científicas, mientras que las restantes son pragmáticas. De acuerdo con van Fraassen, las siguientes metas son meramente pragmáticas: la elegancia matemática, la simplicidad, el amplio alcance, que sea completa en algunos aspectos, que sea capaz de unificar fenómenos dispares y el poder explicativo (*Cfr.* van Fraassen, 1980, p.87). Así, todas las metas de una teoría diferentes de la adecuación empírica son pragmáticas. Lo más importante es que van Fraassen ubica a la explicación científica dentro de la clase de metas pragmáticas de la ciencia, no dentro de la clase de metas epistémicas. Si van Fraassen tuviera la razón, entonces aquí habría un problema, pues he aceptado que la explicación científica es una meta epistémica de la actividad científica. Argumentaré que esta dificultad es meramente aparente.

Al examinar con cuidado las razones por las que van Fraassen sostiene que la explicación es una meta pragmática y no epistémica resulta evidente que incluso la noción de explicación (supuestamente) pragmática de van Fraassen podría ser explicada en los términos de la TVEE. Según él, la explicación científica es una meta independiente de la verdad o la adecuación empírica: “[la explicación] no tiene que ver con la relación entre la teoría y el mundo, sino más bien con el uso y utilidad de la teoría; da razones para preferir la teoría independientemente de cuestiones sobre su verdad” (van Fraassen, 1980, p. 88). Pero al examinar con mayor detalle, resulta que efectivamente la verdad tiene un papel esencial en la explicación del valor epistémico de la explicación científica, lo cual permitiría explicar el valor de la explicación científica en términos de la TVEE. Veamos cómo se logra esto. El propio van Fraassen

reconoce que la búsqueda de la explicación es una empresa racional. Según él mismo, una teoría debe satisfacer ciertas metas epistémicas (debe ser verdadera o adecuada empíricamente, dependiendo de si se defiende un realismo o un empirismo constructivo) para poder suministrar explicaciones satisfactorias:

Los méritos epistémicos que una teoría puede o debe tener para figurar en buenas explicaciones no son *sui generis*; son precisamente los méritos que tiene al ser empíricamente adecuada, o al tener fuerza empírica significativa, etc. [...] Al buscar la explicación buscamos a fortiori aquellos méritos más básicos, esto es lo que hace valiosa a la búsqueda de la explicación para la empresa científica como tal. (van Fraassen, 1980, p. 88).

Esto significa que una buena explicación sólo puede surgir de una buena teoría, esto es, de una teoría que efectivamente tenga algunas virtudes genuinamente epistémicas. Van Fraassen insiste en que el hecho de que una teoría sea explicativa no dice nada sobre la relación de la teoría con el mundo y que eso es una razón suficiente para no considerar a la explicación como una meta epistémica de la ciencia. Él tiene todo el derecho a adoptar esta clasificación; sin embargo también es cierto que de acuerdo con el veritismo que se ha adoptado en esta investigación, incluso adoptando la elucidación de la explicación científica que presenta van Fraassen resultaría que la explicación es *epistémicamente* valiosa, pues podríamos explicar su valor en términos de la TVEE. Hemos visto que el propio van Fraassen concede que las buenas explicaciones *dependen* de teorías verdaderas, si somos realistas, o de teorías adecuadas empíricamente, si somos empiristas constructivos. Así, el valor de las explicaciones depende de un modo esencial de la verdad (o adecuación empírica) de las teorías que se usan para dar las explicaciones. Por lo tanto, desde el punto de vista veritista que he adoptado aquí, incluso el valor de la explicación científica (entendida al modo de van Fraassen) es también explicable en

términos de la TVEE, pues la explicación científica ejemplifica la relación de dependencia que se definió anteriormente, esto es el valor de una buena explicación depende de que la teoría que se usa para explicar un fenómeno tenga algún valor epistémico previo, así sea la verdad o la adecuación empírica.

Pero hay que recordar que según la definición de la relación de dependencia, el valor de la unidad de evaluación debe depender de la creencia verdadera o de otro bien epistémico explicable en términos de la TVEE. van Fraassen, siendo un empirista constructivo, insistiría en que el valor de la explicación científica depende de la adecuación empírica de las teorías, *no* de su verdad. Para lograr dar cuenta de la propuesta de van Fraassen usando el marco veritista hace falta explicar también cómo el valor la adecuación empírica podría explicarse en términos de la TVEE. Esta tarea, sin embargo, no resulta difícil. Hay que recordar que según van Fraassen, la adecuación empírica es la *verdad* sobre los fenómenos observables, es decir es un tipo particular de verdad. Así, desde el punto de vista del empirista constructivo, el valor de la explicación científica depende de que la teoría se apoye en un tipo particular de verdades. Así, incluso adoptando el marco del empirismo constructivo, el valor de la explicación puede explicarse en términos de la TVEE.

Ahora bien, la impresión de que el valor epistémico de la explicación científica (según es entendida por van Fraassen) es explicable según la TVEE se ve fortalecida cuando se analiza con más detalle la propuesta específica de van Fraassen sobre cómo caracterizar a la explicación científica. En particular, su propuesta sostiene que la forma correcta de toda pregunta que busca explicación es la siguiente: “¿Por qué P , en contraste con X ?”¹³. Además sostiene que toda pregunta de este tipo tiene una *presuposición* asociada y que ésta está constituida por, al menos dos supuestos particulares: a) que P es un miembro de X y b) que P es

¹³ Donde P refiere, típicamente, a una proposición que describe un fenómeno particular y X refiere a un conjunto de proposiciones que describen fenómenos alternativos.

verdadera y la mayoría de las X s no lo son (*Cfr.* van Fraassen, 1977, p. 149). Un ejemplo ayudará a entender mejor la propuesta: supongamos que buscamos una explicación de por qué Franz Liszt tocó en su primer concierto en Kiev una sonata para piano particular (P), en lugar de tocar cualquier otra pieza de su repertorio (X), como un trío para piano, violín y cello o una sonata para piano distinta de la que de hecho tocó. De acuerdo con la postura de van Fraassen, la pregunta misma presupone a) que la sonata para piano que tocó (P) pertenece a la clase de las piezas musicales de su repertorio (X) y b) que es verdad que tocó P y es falso que tocó las otras piezas de su repertorio. Así, una explicación correcta (es decir una respuesta correcta a la pregunta “¿Por qué tocó esa pieza y no otra?”) consistiría en decir que Liszt se enteró de que Chopin iba a asistir a su concierto y, preocupado por impresionarlo, eligió tocar la pieza técnicamente más complicada de su repertorio, es decir la sonata para piano que de hecho tocó. Volviendo al tema central, lo importante de esta propuesta es que nuevamente la verdad se hace presente de manera evidente y esencial en la parte (b) de la presuposición asociada a toda pregunta que busca explicación; esto es, esa presuposición indica que lo que buscamos es una respuesta verdadera a la pregunta que busca explicación. Es cierto, que la propuesta de van Fraassen es mucho más compleja y detallada de lo que he presentado aquí. Sin embargo, parece que esto basta para mostrar que efectivamente no resulta problemática para el veritismo que se asume en este trabajo.

Ahora sí podemos pasar a analizar con detalle los tipos de entendimiento científico que identifiqué en la sección anterior. Comenzaré con las posturas que relacionan el entendimiento con la explicación científica y con la unificación teórica. Dejaré para el capítulo 3 el análisis del entendimiento científico asociado con el uso de modelos idealizados.

2. Explicación veritista de algunos tipos de entendimiento científico

En este capítulo analizaré con detalle cuatro tipos de entendimiento científico. Tres de ellos tienen que ver con distintas teorías sobre la explicación científica y el restante tiene que ver con la noción de unificación teórica. En todos los casos, investigaré si es posible explicar el valor de estos tipos de entendimiento según la TVEE. Tradicionalmente, los filósofos de la ciencia han

insistido en que las explicaciones científicas nos proporcionan entendimiento del fenómeno explicado. Por esta razón he elegido analizarlas con detalle. Es cierto que esta conexión entre explicación y entendimiento rara vez es abordada explícitamente por los filósofos de la ciencia y, cuando lo hacen, lo hacen de una manera superficial. Por esta razón, cuando presente cada modelo de explicación mostraré cómo los propios autores reconocen que hay una relación importante entre explicación y entendimiento, si bien no la analizan con detalle.

Hay que recordar que, en concordancia con lo que sostiene la TVEE, la metodología consistirá en explicar el valor de los distintos tipos de entendimiento científico apelando a alguna relación adecuada que tienen éstos con la creencia verdadera. En la sección 2.2 presenté las relaciones de las que haré uso en esta investigación para llevar a cabo esta tarea explicativa: 1) dependencia con la verdad 2) conductividad a la verdad y 3) ser altamente probable, dada la evidencia.

2.1 La explicación científica

2.1.1 El modelo nomológico-deductivo e inductivo-estadístico

Este modelo de explicación científica puede ser visto como el modelo clásico. Fue propuesto por Hempel y Oppenheim en los años 50s. De acuerdo con éste hay dos tipos de explicaciones científicas fundamentales: las nomológicas-deductivas (N-D) y las inductivo-estadísticas (I-E). Ambos tipos de explicaciones *son* tipos de argumentos: las explicaciones N-D son argumentos deductivos, mientras que las I-E son argumentos inductivos. Para que una explicación científica (así sea I-E o N-D) sea adecuada, los argumentos deben cumplir otros requisitos. Pero antes de pasar a estos hay que dar algo de terminología. El evento a ser explicado será llamado *explanandum* y la oración que lo describe se llamará *oración explanandum*. Por otro lado, las premisas del argumento explicativo serán llamadas *explanans* y éste deberá incluir al menos

una oración que describe una ley de la naturaleza y podrá incluir oraciones que describen situaciones particulares.

Comencemos con las explicaciones N-D. Para que una explicación científica de este tipo sea satisfactoria es necesario que se cumplan las siguientes condiciones: 1) la oración *explanandum* debe seguirse deductivamente del *explanans* y las oraciones que forman el *explanans* deben ser verdaderas. En otras palabras, la explicación N-D debe tomar la forma de un argumento deductivo sólido (i. e. que sea válido y que tenga premisas verdaderas) en el que el *explanandum* se sigue lógicamente del *explanans*. 2) El *explanans* debe estar constituido por al menos una ley de la naturaleza y esta premisa debe ser esencial para la explicación, en el sentido de que la derivación no sería válida si se eliminara esta ley de la naturaleza.

Un ejemplo típico de este tipo de explicaciones es el siguiente:

1. Todos los metales son buenos conductores de calor y electricidad.
2. El oro es un metal.
3. El anillo de compromiso de mi hermana está hecho de oro.
4. Por lo tanto, el anillo de mi hermana es un buen conductor de calor y electricidad.

El modelo N-D descansa en gran medida en el concepto de “ley de la naturaleza”, de modo que si se desea dar una explicación satisfactoria de este modelo hay que dar una elucidación clara de este concepto. Según las concepciones más tradicionales, las leyes de la naturaleza son oraciones de forma universal que describen ciertas regularidades de la naturaleza que no aceptan excepciones. Sin embargo, no todas las oraciones de forma universal son leyes de la naturaleza, pues algunas son meras generalizaciones accidentales,¹⁴ por ejemplo, la oración “todas las rocas en esta caja contienen hierro”. Sin embargo, como bien señala James

¹⁴ El propio Hempel señala algunos modos en que se podría dar un criterio para distinguir a las leyes genuinas de las generalizaciones accidentales en el capítulo 5 de su *Philosophy of Natural Science* (1966).

Woodward, a pesar de que se han propuesto varios criterios para distinguir a las leyes genuinas de las meras generalizaciones accidentales, ninguna ha ganado aceptación general (Cfr. Woodward, 2011, pp. 7-10). Aquí no profundizaré en este problema, pues no es esencial para el objetivo que persigo. Basta con caracterizar la explicación N-D y señalar que hace falta decir algo más para dar una elucidación clara del concepto de ley de la naturaleza, tan relevante para este modelo explicativo.

Ya dijimos que, según Hempel, toda explicación científica satisfactoria puede formularse en términos de argumentos. Sin embargo, a diferencia del modelo N-D, el modelo I-E propone una explicación en la que el argumento en cuestión no es deductivo sino inductivo. Hay, sin embargo, tres condiciones extras que debe cumplir un argumento inductivo para ser una explicación I-E satisfactoria: 1) las premisas que forman el *explanans* deben ser verdaderas y debe incluir al menos una ley probabilista, 2) si bien la conclusión no se sigue deductivamente de las premisas, sí debe seguirse con una alta probabilidad, 3) finalmente, la explicación debe cumplir con el requisito de máxima especificidad. Enseguida explicaré con mayor detalle cada condición.

Además de establecer que las premisas que forman el *explanans* deben ser verdaderas, la condición (1) establece que toda explicación I-E debe incorporar en su *explanans* al menos una “ley de forma estadística básica”, que son del siguiente tipo

$$p(G,F) = r$$

Esto quiere decir que “a largo plazo la proporción de los casos de F que también son G es aproximadamente r ” (Hempel, 1965, p. 376). Es importante notar que (1) no exige que las leyes probabilistas utilizadas en una explicación I-E establezcan que la probabilidad de que G sea F sea alta, es decir no se exige que el valor de r sea alto. Lo único que se exige es que el suceso a explicar, la conclusión del argumento, se infiera con un alto grado de probabilidad, en particular esta exigencia se encuentra explícita en la condición (2). Por ejemplo, supongamos

que hay una regla estadística que establece que la probabilidad de obtener un seis al lanzar un dado es de $1/6$. Esta regla no establece que la probabilidad de sacar un seis sea alta, de hecho en cada lanzamiento particular es más probable que salga un número distinto del seis. Sin embargo, podemos utilizar esta regla en una explicación I-E, siempre y cuando la conclusión del argumento sea altamente probable, dadas las premisas. Por ejemplo:

1. La probabilidad de sacar un seis al lanzar un dado es de $1/6$. Es decir $P(S) = 1/6$.
2. *A* lanzó el dado doce veces.
3. Por lo tanto, es altamente probable que *A* haya sacado un seis.

Finalmente, la condición (3) sostiene que las explicaciones probabilistas deben de cumplir el requisito de “máxima especificidad”, que sostiene lo siguiente: “cuando se elige la clase a la que pertenece el caso individual que queremos explicar, [...] no debemos saber cómo dividirla en sub-conjuntos en los que la probabilidad del hecho a ser explicado difiera de su probabilidad en la clase completa” (Salmon, 1984, p. 29). Es decir, debemos ubicar al fenómeno en una clase tal que (dado nuestro conocimiento) no podamos subdividirla, a su vez, en otras clases con una probabilidad distinta de la clase inicial. Esto garantiza que todos los elementos relevantes estarán incluidos en la subclase que elegimos, pues hemos dado con la subclase más específica, dado nuestro conocimiento de la situación. Podemos entender por qué se introduce esta condición si consideramos un ejemplo que presenta el propio Hempel:

Una explicación de la recuperación de Jones basada en la información de que Jones tuvo una infección por estreptococo y que fue tratado con penicilina, y que la probabilidad estadística de recuperación en tales casos es alta es inaceptable si *K* [el conjunto de las oraciones aceptadas en esa situación] incluye la información adicional de que el estreptococo de Jones

era resistente a la penicilina, o que Jones era un octogenario con un corazón débil, y que en esa clase de referencia la probabilidad de recuperación es pequeña (Hempel, 1965, p. 398).

La razón por la que se rechaza que una explicación como ésta sea buena consiste precisamente en que viola el requisito de máxima especificidad. El error consiste en incluir el suceso que queremos explicar (i.e. la recuperación de Jones) en la clase de personas con una infección por estreptococo tratadas con penicilina, cuando de hecho tenemos información adicional (que la sepa que infectó a Jones era resistente a la penicilina o que él tiene más de ochenta años) que es relevante para clasificar al suceso a explicar en una subclase mucho más específica, es decir una subclase cuya probabilidad de recuperación difiere drásticamente de la clase original.

Ahora que hemos expuesto de manera sucinta los modelos N-D e I-E de explicación científica y antes de pasar a ver de qué manera podría explicar el veritista el valor epistémico que emana de ellos, me gustaría señalar que el propio Hempel creía que su modelo de explicación tenía una relación cercana con el entendimiento científico. Según él las explicaciones científicas nos proporcionan entendimiento:

...una explicación N-D responde a la pregunta “¿*Por qué* el fenómeno-*explanandum* ocurrió?” al mostrar que el fenómeno resultó de ciertas circunstancias particulares, especificadas en C_1, C_2, \dots, C_k , de acuerdo con las leyes L_1, L_2, \dots, L_r . Al indicar esto el argumento muestra que, dadas las circunstancias particulares y las leyes en cuestión, la ocurrencia del fenómeno *era esperable*; y es en este sentido que la explicación nos permite *entender por qué* el fenómeno ocurrió (Hempel, 1965, p. 337).

En otras palabras, entendemos por qué ocurrió un fenómeno cuando tenemos una explicación científica adecuada del mismo, tal que nos lleva a esperar la ocurrencia del

fenómeno en cuestión. Ahora bien, de acuerdo con Woodward, las explicaciones I-E también nos proporcionan entendimiento a pesar de que no proporcionan la misma expectativa que las explicaciones N-D. Las explicaciones I-E involucran una generalización de la idea central en la cita anterior: “a pesar de que una explicación I-E no muestra que el fenómeno-explanandum era esperable con certeza, hace la siguiente cosa buena: muestra que el fenómeno-explanandum es al menos esperable con alta probabilidad y de este modo nos proporciona entendimiento” (Woodward, 2011, p. 12). De este modo, hay un vínculo entre la explicación científica y el entendimiento científico.

Pasemos ahora a revisar de qué manera podría explicarse en términos veritistas el valor epistémico que tienen las explicaciones científicas según el modelo N-D e I-E. Comencemos con el modelo N-D. Aquí encontramos un sentido muy claro en el que se cumple la relación de dependencia con la creencia verdadera. El propio Hempel dice explícitamente que una condición que debe cumplir toda explicación N-D satisfactoria es que “las oraciones que constituyen el explanans deben ser *verdaderas*” (Hempel y Oppenheim, 1948, p. 137).¹⁵ Esto significa que no basta con que los argumentos que proporcionan explicación científica sean válidos, sino que además deben ser sólidos. De acuerdo con la teoría de Hempel, uno entiende un fenómeno cuando tiene una explicación N-D cuyo *explanandum* describe el fenómeno en cuestión. Así, entender un fenómeno es razonar del modo que prescribe una explicación N-D. Dada la naturaleza de la explicación N-D se sigue que la actividad de entender un fenómeno mediante una explicación de este tipo ejemplifica la relación de dependencia con la creencia verdadera. Pues en una explicación N-D las premisas que forman el *explanans* son verdaderas. Por lo tanto, si entender un fenómeno consiste en razonar como prescribe una explicación N-D, entonces, en última instancia, entender un fenómeno depende de tener ciertas creencias verdaderas, a saber creencias en las proposiciones que aparecen en el *explanans*. Pero hay más.

¹⁵ Las cursivas son mías.

Recordemos que un requisito para que una explicación científica sea aceptable es que el *explanandum* se siga mediante una inferencia válida del *explanans*. Combinando las propiedades de validez y solidez de las explicaciones científicas N-D, tenemos que *necesariamente* el *explanandum* será verdadero. Por lo tanto, el valor epistémico de entender un fenómeno apelando a una explicación N-D no sólo proviene de que el proceso de razonamiento dependa de creencias verdaderas; además, razonar de este modo tiene una relación de conductividad con la creencia verdadera: adoptar como creencias las conclusiones de un argumento N-D nos lleva a formar creencias *verdaderas* de un modo fiable. De modo que entender un fenómeno mediante una explicación N-D también ejemplifica la relación de conductividad a la verdad. Así el valor de este tipo de entendimiento científico depende del valor de la creencia verdadera en la medida que ejemplifica estas dos relaciones.

Sin embargo, se podría argumentar que el requisito de verdad de las premisas que componen el *explanans* es muy fuerte. Por ejemplo, se podría debilitar el requisito, eliminando la exigencia de que las premisas sean verdaderas y sosteniendo, en su lugar, que basta con que éstas estén epistémicamente justificadas. El modelo de Hempel sí considera que las premisas que forman el *explanans* deben estar justificadas, pero además añade la exigencia de que éstas sean verdaderas. Me parece que este requisito podría resultar demasiado exigente y, por lo tanto, sugiero la contracción del requisito que señala que basta con que las premisas estén justificadas. El propio Hempel considera esa alternativa y la descarta argumentando que tendría la siguiente consecuencia negativa: una explicación científica podría ser buena en un momento histórico determinado y, posteriormente, dejar de serlo (*Cfr.* Hempel y Oppenheim, 1948, p. 137). Sin embargo, no es claro que todos debamos compartir su intuición de que esta consecuencia es indeseable, sobre todo tomando en cuenta que, al echar un vistazo a la historia de la ciencia, efectivamente nos encontramos con explicaciones que la propia comunidad

científica consideraba aceptables en un momento y que en un momento posterior dejó de hacerlo.

Parece, sin embargo, que incluso adoptando el requisito más débil, que sólo exige justificación (y no verdad) para las premisas del *explanans*, el veritista puede explicar en sus términos el valor de entender un fenómeno apelando a una explicación N-D. Asumo aquí una definición de justificación como la que propone William Alston: “estar justificado en creer que p es creer que p de una manera tal que es al menos muy probable que la creencia de uno sea verdadera” (Alston, 1989, p. 2). Asumo esta definición no porque haga la tarea más sencilla para el veritista, sino porque es ampliamente aceptada por los epistemólogos contemporáneos.¹⁶ Esta definición de justificación epistémica implica que formar creencias justificadas es un método fiable de formación de creencias, es decir un método que arroja una proporción grande de creencias verdaderas: si una condición para que una creencia esté justificada es su alta probabilidad, entonces se sigue que formar creencias justificadas es un método conducente a la verdad, por una mera cuestión probabilística. Así entendida, la creencia verdadera tiene un papel esencial en la explicación del valor de la justificación epistémica: la justificación epistémica es valiosa porque ejemplifica la relación de conductividad a la verdad con la creencia verdadera. De manera más precisa: aceptar creencias sólo si están justificadas epistémicamente es un método fiable de formación de creencias. De modo que el valor de la justificación epistémica también es explicable en los términos de la TVEE.¹⁷ Así, incluso en esta formulación alternativa del requisito de Hempel para las premisas del *explanandum* (que

¹⁶ Sólo por presentar un ejemplo, James Pryor defiende una definición de justificación epistémica similar. Él dice que “tú tienes justificación para creer P si y sólo si estás en una posición en la que sería apropiado epistémicamente para ti creer P , una posición en la que P es probablemente verdadera para ti” (Pryor, 2005, p. 6).

¹⁷ Hay, sin embargo, algunas concepciones de la justificación epistémica que no se ajustan a la definición que presenté aquí. En particular, algunas teorías internistas de la justificación desafían el supuesto de que el requisito de fiabilidad sea necesario para dar una definición satisfactoria del concepto de justificación epistémica. Más adelante, en la sección 3.6, aparecerá nuevamente y de un modo más relevante la cuestión sobre la justificación epistémica. En ese punto explicaré con mayor detalle qué respuestas podría dar el veritista a quienes defienden una concepción de la justificación de este tipo.

requiere mera justificación epistémica y no verdad) es posible explicar el valor de entender un fenómeno mediante una explicación N-D en términos de la TVEE: entender de este modo un fenómeno depende esencialmente de la justificación epistémica a favor de las premisas que forman el *explanans* y, a su vez, el valor de la justificación epistémica es explicable en términos de la TVEE.

Pasemos ahora a examinar el modelo I-E. Aquí, al igual que con el modelo N-D, se exige que las premisas que forman el *explanans* sean *verdaderas*.¹⁸ Por lo tanto, entender un fenómeno apelando a una explicación I-E ejemplifica también la relación de dependencia con la creencia verdadera. Sin embargo, hay otro sentido en el que la explicación I-E está relacionada íntimamente con la verdad. Recordemos que uno de los requisitos que debe cumplir un argumento inductivo para contar como una buena explicación I-E es el requisito de alta probabilidad, que exige que el argumento inductivo tenga como resultado que el fenómeno-*explanandum* sea altamente probable, dadas las premisas que forman el *explanans*. Si uno entiende un fenómeno apelando a una explicación I-E, entonces razona del modo que prescribe ese argumento. Dado el requisito de alta probabilidad, de aquí se sigue que razonar de este modo arroja creencias que son altamente probables, dada la evidencia que se presenta en las premisas del argumento inductivo-estadístico. Así, entender un fenómeno apelando a una explicación I-E tiene como resultado la formación de creencias que ejemplifican la relación de ser altamente probables, dada la evidencia que se presenta en las premisas.

Tenemos por lo tanto, el siguiente resultado: el valor epistémico del tipo de entendimiento científico que surge del uso de explicaciones I-E y N-D es explicable en términos de la TVEE, pues los tipos de entendimiento que surgen de ambos tipos de explicación ejemplifican las tres relaciones con la creencia verdadera que presenté en la sección (1.2). Además, si las

¹⁸ De nuevo, quizás este requisito resulte demasiado exigente. Sin embargo, una modificación que exija sólo la justificación de las premisas también resultaría explicable en términos de la TVEE como se mostró en el caso de la explicación N-D.

explicaciones de estos tipos que se usan para entender un fenómeno no ejemplificaran estas relaciones con la creencia verdad, no consideraríamos que son epistémicamente valiosas. Imaginemos que tenemos explicaciones de fenómenos cuyos *explanans* están constituidos por proposiciones falsas e injustificadas, que no son conducentes a la verdad y que no arrojan como conclusiones proposiciones con alta probabilidad, dada la evidencia. Claramente no consideraríamos que las explicaciones de estos tipos son epistémicamente valiosas.

2.1.2 La explicación por relevancia estadística

A continuación revisaré un tipo de entendimiento científico que surge de una teoría de la explicación distinta a la que propone Hempel: la teoría que se centra en la relevancia estadística, defendida principalmente por Wesley C. Salmon. De acuerdo con él, el entendimiento científico está estrechamente relacionado con la explicación científica. Cuando nos involucramos en la ciencia, sostiene Salmon, no sólo buscamos saber que suceden ciertos fenómenos, sino que también deseamos entender por qué suceden. Esto parece sugerir que el entendimiento científico es algo diferente (más complejo) que las meras descripciones de fenómenos. Aparentemente para responder a la siguiente pregunta: “¿qué, además del conocimiento descriptivo del mundo, se requiere para alcanzar entendimiento?” (Salmon, 1984, p. 9), es necesario investigar qué es una explicación científica adecuada. Todo parece indicar que, según él, aquello que vuelve a un conjunto de información una explicación de un fenómeno es lo mismo que nos da entendimiento de ese fenómeno.

Ahora bien, Salmon cree que su teoría de relevancia estadística es mejor que la teoría de Hempel en la medida en que supera dos problemas que aquella no podía solucionar. El primer problema consiste en que hay argumentos que satisfacen todos los requisitos que presenta Hempel y que, sin embargo, no constituyen una explicación científica satisfactoria. El ejemplo que presenta Salmon para convencernos de esto es uno en el que se hace uso de información

irrelevante en la explicación. Supongamos que tenemos una muestra de sal de mesa que se disuelve cuando la mezclamos con agua. Éste es el fenómeno-*explanandum*. Al elaborar una explicación se presentan las siguientes premisas: a) Alguien lanzó un hechizo sobre la sal para que se disolviera y b) todas las muestras de sal de mesa hechizadas se disuelven en agua.¹⁹ Y a partir de estas premisas se deriva deductivamente el fenómeno-*explanandum*. (Cfr., Salmon, 1984, p. 31) Este argumento cumple con los requisitos que establece Hempel y, sin embargo, claramente no es una explicación científica satisfactoria. Hay una variedad de diagnósticos que explican cuál es la fuente de este error en la teoría de Hempel. Por ejemplo, Woodward argumenta que el error consiste en que una explicación satisfactoria debe citar las causas del fenómeno-*explanandum* y que el ejemplo que presenta Salmon resulta en una explicación insatisfactoria porque no cita las causas del fenómeno a explicar. Claramente, Woodward favorece este diagnóstico pues él mismo está interesado en defender una teoría causal de la explicación científica. Por otro lado, Salmon insiste en que el error en la teoría de Hempel que permite la existencia de estos ejemplos problemáticos consiste en que su requisito de máxima especificidad no impide que se incluya información irrelevante en el *explanans*.²⁰ Por esta razón Salmon favorece un modelo en el que la relevancia de la información que se presenta en el *explanans* con relación al fenómeno-*explanandum* es central. Sin embargo, independientemente de cuál sea el diagnóstico de qué provoca estos problemas en la teoría de Hempel, es claro que superar esta dificultad es un desiderátum para toda teoría de la explicación científica.

¹⁹ Hay que recordar que en la sección anterior llamé la atención sobre el hecho de que una presentación satisfactoria del modelo N-D debe incluir una noción precisa de “ley de la naturaleza”. Es claro que en el ejemplo que presenta Salmon, la generalización de la premisa (b) no sería considerada una ley de la naturaleza por muchos. Un defensor del modelo N-D podría argumentar que el ejemplo que presenta Salmon no cumple los requisitos que propone Hempel, pues la generalización de la premisa (b) no es una ley de la naturaleza. Como en esta investigación no busco determinar qué modelo explicativo es más atractivo no profundizaré en esta cuestión

²⁰ Salmon sostiene que “El requisito de máxima especificidad de Hempel garantiza que *todos* los hechos relevantes conocidos deben incluirse en una explicación científica adecuada, pero no hay un requisito de que sólo los hechos relevantes se incluyan” (Salmon, 1984, p.31)

El segundo problema que enfrenta la teoría de Hempel tiene que ver con el requisito de alta probabilidad y consiste en que, contrario a lo que sugiere el modelo I-E de Hempel, claramente hay explicaciones científicas adecuadas cuyo fenómeno-*explanandum* *no* es altamente probable dado el *explanans*. Consideremos el siguiente ejemplo: la explicación correcta de por qué un sujeto desarrolló paresia consiste en señalar que ese sujeto está infectado de sífilis y que la enfermedad ha progresado hasta la tercera etapa sin ser tratado con penicilina. La paresia es una forma de sífilis terciaria y nunca ocurre más que en sifilíticos. Así, el hecho de que un sujeto tenga sífilis latente no tratada explica por qué desarrolló paresia. Sin embargo, resulta que la probabilidad de desarrollar paresia en las víctimas de sífilis no tratada *no* es alta. De hecho es más probable no desarrollarla (*Cfr.* Salmon, 1984, pp. 31, 32). Esto muestra que “el requisito de alta probabilidad no constituye una condición necesaria para las explicaciones estadísticas legítimas” (Salmon, 1984, p. 32). Así, el segundo desiderátum del modelo de Salmon será rechazar el requisito de alta probabilidad, para poder dar cuenta de casos como el de la paresia.

El modelo de explicación que propone Salmon hace uso de la noción relevancia estadística y no de la alta probabilidad para caracterizar las explicaciones científicas adecuadas. Por esta razón, su modelo suele conocerse como el modelo de *relevancia estadística* (R-E, en adelante). De acuerdo con Salmon, “*C* es un factor estadísticamente relevante para la ocurrencia de *B* bajo las circunstancias *A* si y sólo si

$$P(B|A.C) \neq P(B|A)” \text{ (Salmon, 1984, pp. 32, 33).}$$

Será útil recordar el ejemplo de Hempel sobre el estreptococo para aclarar esta noción. Decimos que el hecho de que el tipo de estreptococo que infectó a Jones sea resistente a la penicilina (*C*) es un factor estadísticamente relevante para su recuperación (*B*), asumiendo que toma penicilina (*A*), si la probabilidad de que se recupere, dado que toma penicilina y tiene una infección resistente a la penicilina ($P(B|A.C)$) es diferente de la probabilidad de que se recupere, dado que toma penicilina ($P(B|A)$), es decir $P(B|A.C) \neq P(B|A)$.

Una vez equipados con la noción de relevancia estadística podemos mostrar claramente en qué consiste la teoría de la explicación científica de Salmon. De acuerdo con él, dar una explicación adecuada de un fenómeno involucra hacer referencia a un factor estadísticamente relevante que posee el fenómeno-*explanandum*, con respecto a una clase inicial de referencia. Ilustrando este punto con el ejemplo de Jones, diríamos que una explicación de que su recuperación fue lenta consiste en señalar que su infección era resistente a la penicilina y que éste es un factor estadísticamente relevante para su recuperación, con respecto a la clase de referencia de todos los casos de infección por estreptococo tratada con penicilina. De acuerdo con Salmon, la manera correcta de formular la pregunta que busca una explicación es ésta: “¿Por qué x , que es A , es también B ?” y la respuesta correcta debe ser una de este estilo: “porque x es también C , donde C es *relevante* para ser B dentro de A ” (Cfr. Salmon, 1984, pp. 34, 35). En el caso de la infección por estreptococo, la pregunta sería: ¿Por qué Jones (que está infectado con estreptococo), a pesar de ser tratado con penicilina, no se recuperó rápido? Y la respuesta sería: la infección de Jones es resistente a la penicilina, y el hecho de que sea resistente es relevante para su recuperación rápida, incluso cuando haya sido tratado con penicilina.

Llamemos *probabilidad anterior* a $P(B|A)$ y llamemos *probabilidad posterior* a $P(B|A.C)$. Una explicación satisfactoria, según el modelo de Salmon, consiste en señalar que la probabilidad anterior de la ocurrencia de un evento es distinta de su probabilidad posterior. Si ilustramos esto con el ejemplo del estreptococo quedará más claro. Hay que comenzar estableciendo la probabilidad anterior de recuperación para todos los casos de infección con estreptococo. A continuación se consideran cuáles son los factores relevantes que pueden influir en el proceso de recuperación de la infección. Se divide a la clase inicial (la clase de todos los infectados con estreptococo) de acuerdo con estos factores relevantes en células (sub-clases) y, finalmente se asigna una probabilidad a cada una de estas células. En el caso del

estreptococo, hay dos factores relevantes para la recuperación: el tratamiento con penicilina y la existencia de sepas resistentes a la penicilina. Así, la clase original de referencia se divide en cuatro células: (1) Infección por bacteria no resistente, penicilina administrada; (2) infección por bacteria no resistente, sin penicilina administrada; (3) infección por bacteria resistente, penicilina administrada y (4) infección por bacteria resistente, sin penicilina administrada. Supongamos que el caso de Jones pertenece a la célula (3). En este caso tenemos una explicación adecuada de su recuperación lenta, no porque la probabilidad de que se recupere lentamente sea alta, sino porque la probabilidad de recuperación rápida de los sujetos en la célula (3) difiere significativamente de la probabilidad de recuperación en la clase inicial de referencia (*Cfr.* Salmon, 1984, p. 33). Señalar el hecho de que el caso de Jones pertenece a una clase que tiene una característica estadísticamente relevante con respecto a la clase inicial ya es, de acuerdo con Salmon, una explicación exitosa de por qué Jones se recuperó lentamente de la infección.

Ahora bien, es necesario explicar de qué manera el modelo R-E es capaz de superar las dificultades que enfrentaba el modelo N-D/I-E, pues hay que recordar que Salmon asume esta tarea como un desiderátum de su teoría. La primera dificultad es la de la irrelevancia. Volvamos al ejemplo de la sal hechizada. No podemos apelar al “hecho” de que la sal ha sido hechizada para explicar la disolución de una muestra particular en agua porque el “hecho” de que la sal ha sido hechizada es estadísticamente irrelevante para su solubilidad. Es decir, la probabilidad de que cualquier muestra de sal hechizada sea soluble es igual a la probabilidad de que cualquier muestra de sal sea soluble. Así, la teoría de Salmon garantiza que toda información que se utilice para explicar un fenómeno sea relevante para su ocurrencias. En segundo lugar, tenemos el problema de las explicaciones con baja probabilidad. Este problema claramente es superado por el modelo R-E, pues, a diferencia del modelo de Hempel, el modelo R-E no requiere que el fenómeno-*explanandum* sea altamente probable. Retomemos el ejemplo

de la paresia. Para explicar por qué un sujeto particular desarrolló paresia basta con que señalemos que en el caso particular estaba presente una característica estadísticamente relevante. En otras palabras, basta con señalar a un factor estadísticamente relevante para tener una explicación satisfactoria, no hace falta que haya un argumento que vuelva altamente probable al fenómeno-*explanandum*.

Aunque sea meramente un punto tangencial dados los intereses de esta investigación, es importante señalar que de este modelo se sigue una consecuencia bastante anti-intuitiva: un mismo *explanans* puede explicar dos hechos contradictorios. Retomemos el caso de Jones y el estreptococo. En el caso de la recuperación lenta de Jones, el *explanans* consiste en señalar que el caso de Jones pertenece a la célula (3), es decir la clase a la que pertenecen los casos de infección por una cepa resistente a la penicilina, y en que las características de esa célula son relevantes estadísticamente con respecto a la clase de referencia original. Sin embargo, asumamos que Jones, en la misma situación, se recupera rápidamente de su infección. En tal caso, tendríamos que explicar el fenómeno apelando al mismo *explanans*. Es decir, el hecho de que el caso de Jones pertenece a la célula (3) permite explicar tanto su recuperación rápida como su recuperación lenta. Salmon, sin embargo, considera que hay que aceptar esta consecuencia y que de hecho no representa un problema.²¹ Un *explanans* es correcto siempre y cuando invoque a las probabilidades correctas y cite todos y sólo los factores relevantes para el caso particular (Cfr. Woodward, 2011, pp. 26-27). Por su parte, Woodward sugiere que para aliviar este problema el defensor del modelo R-E podría apelar a que el fenómeno-*explanandum* de toda explicación R-E no es un evento particular sino una generalidad

²¹ Quizás la reacción más natural ante esa dificultad sea intentar subdividir la célula (3), a partir de nuevos factores estadísticamente relevantes. Como ya vimos, la postura de Salmon consiste en tomar al toro por los cuernos e intentar argumentar a favor de que esta consecuencia no es tan anti intuitiva después de todo. Aquí no profundizaré sobre cuál es la postura correcta ante esta problemática, pues no es esencial para la finalidad que persigo con esta discusión.

estadística. No diré aquí más sobre este problema, pues mi interés central no está en este asunto.

A continuación analizaré de qué manera una teoría veritista podría dar cuenta del valor epistémico del entendimiento que surge de una explicación R-E. De acuerdo con esta teoría, entender un fenómeno consiste en tener una respuesta adecuada a la pregunta ¿Por qué x , que es A , es también B ? A su vez, una respuesta adecuada a esta pregunta consiste en señalar que x tiene alguna característica adicional C que es estadísticamente relevante para que x sea B . Para que una explicación sea adecuada es necesario que las proposiciones probabilísticas y las proposiciones particulares que se citan sean verdaderas o que por lo menos estén epistémicamente justificadas. Retomando el caso de la infección de Jones, una explicación adecuada de su lenta recuperación consiste, al menos, en citar el hecho de que Jones tiene una infección resistente a la penicilina y el hecho de que este factor es estadísticamente relevante para la recuperación de esta enfermedad. Así, para tener entendimiento de un fenómeno (por ejemplo, de la lenta recuperación de Jones) uno debe de tener ciertas creencias *verdaderas* sobre el tema que entiende, por ejemplo que Jones tiene una infección resistente a la penicilina y que éste es un factor estadísticamente relevante para que se dé el fenómeno. Por lo tanto, el entendimiento que surge de las explicaciones R-E *depende* de ciertas creencias verdaderas. Así, su valor puede explicarse en términos de la TVEE mediante la relación de dependencia, que sostiene que una unidad de evaluación ejemplifica la relación de *dependencia* con la creencia verdadera cuando su valor depende de que otras creencias (u otras actitudes dóxicas) sean verdaderas o bien tengan un valor epistémico distinto de la verdad que sea explicable en términos de la TVEE. También en este caso si el tipo de entendimiento que surge de las explicaciones R-E no ejemplificara esta relación de dependencia con la creencia verdadera, no consideraríamos que es epistémicamente valioso. Supongamos que explicamos la lenta recuperación de Jones apelando a información falsa sobre el tipo de infección que tiene y sobre

si esto es estadísticamente relevante para su recuperación. No consideraríamos que el “entendimiento” que surge de una explicación defectuosa de este tipo sea epistémicamente valioso.

2.1.3 La explicación causal

A continuación analizaré el entendimiento científico que surge del tipo de explicación causal que defiende James Woodward en su libro *Making Things Happen* (2003). En general, los modelos causales de explicación insisten en la idea de que dar una explicación de un evento consiste en mostrar cómo encaja en una cadena causal o en citar su causa. Woodward sigue de cerca esta idea y le da un contenido específico en su teoría contrafáctica de la explicación causal. Quizás sea importante señalar que su teoría pretende dar una elucidación satisfactoria únicamente de la explicación causal. Es decir, no pretende sostener que toda explicación satisfactoria sea de este tipo, en su lugar propone a esta teoría como la elucidación correcta de sólo un tipo particular de explicación: la explicación causal.

La teoría de Woodward es llamada “contrafáctica” porque descansa en la idea de que las afirmaciones causales pueden analizarse en términos de condicionales contrafácticos. En última instancia su propuesta consiste en sostener que “la explicación es una cuestión de exhibir patrones sistemáticos de dependencia contrafáctica” (Woodward, 2003, p. 191). Hay que explicar cuál es el razonamiento que lleva a Woodward a sostener esta teoría.

Mediante el análisis de casos concretos de explicación científica satisfactoria²², Woodward extrae el desiderátum de que una explicación científica exitosa debe servir para responder a preguntas del tipo “¿Qué hubiera ocurrido si las cosas hubieran sido distintas?”; Woodward

²² Woodward presenta un caso en el que se explica (mediante la ley de Coulomb) por qué la magnitud de la intensidad eléctrica en un sistema concreto obedece alguna generalización matemática. También presenta un caso en el que se explica por qué un monopolio que absorbe una industria previamente competitiva aumentará los precios y reducirá la oferta.

llama a éstas “preguntas-w”. Una explicación exitosa de un fenómeno particular, afirma Woodward, es una que da una respuesta satisfactoria a este tipo de preguntas proveyendo información sobre las causas del fenómeno-*explanandum*. Woodward se compromete con la tesis de que las afirmaciones causales permiten inferir (junto con algunas premisas adicionales) cierta información contrafáctica relevante para la explicación causal. Después de todo, el modelo de Woodward liga de manera cercana la relevancia explicativa con el suministro de un tipo de información contrafáctica.

Retomemos el ejemplo de la paresia presentado cuando se analizó la teoría de Salmon. De acuerdo con Woodward una buena explicación de, por ejemplo, la paresia de Jones es la siguiente: (a) “la sífilis latente de Jones causó su paresia”. Uno podría preguntarse ¿cuál es la información contrafáctica asociada a esta afirmación causal que permite responder a preguntas-w? De acuerdo con Woodward, una afirmación singular causal como ésta, de la forma “*c* causó *e*”, implica el siguiente contrafáctico: si *c* no hubiera ocurrido, entonces *e* no habría ocurrido. Por lo tanto, podemos derivar el siguiente contrafáctico asociado con (a): (b) “Si Jones no hubiera contraído sífilis, no habría desarrollado paresia”. Tenemos entonces que la explicación (a) sobre la paresia de Jones es satisfactoria pues ofrece la información contrafáctica presente en (b) y, además, permite responder a preguntas-w. En particular nos permite responder qué habría sucedido si Jones no hubiera contraído sífilis (*Cfr.* Woodward, 2003, p. 211). Es cierto que este ejemplo es bastante simple y proporciona una explicación poco profunda. Efectivamente, en caso de buscar una explicación más detallada quizás se deberían proporcionar información contrafáctica más fina, que tenga que ver, por ejemplo, con los procesos fisiológicos y neurológicos asociados con la sífilis y la paresia.

Antes de pasar a un análisis más detallado de esta propuesta quizás sea útil mostrar cómo el propio Woodward sostiene que su modelo explicativo proporciona entendimiento científico del mundo. Según el autor, hay un punto general sobre la epistemología de la explicación que

debe respetar toda teoría de la explicación, no únicamente la que él presenta: “una restricción plausible sobre qué es una explicación consiste en que ésta debe ser algo que proporciona entendimiento. Decir que cierta información es 'parte' de una explicación o que contribuye a su importe explicativo es decir que esta información contribuye al entendimiento que provee la explicación” (Woodward, 2003, p. 179). En otras palabras, Woodward afirma que es un desiderátum para toda teoría sobre la explicación que nos diga cómo tener una explicación de un fenómeno también nos da entendimiento de éste. Su propia teoría debe satisfacer esta exigencia si va a ser considerada como una teoría aceptable de la explicación.

Será útil presentar de un modo más preciso la propuesta de Woodward para poder analizar si el valor epistémico del entendimiento que surge de este tipo de explicación es explicable de acuerdo con la TVEE. Woodward presenta así su teoría:

Una explicación relaciona dos componentes, un *explanans* y un *explanandum*; en particular, exhibe cómo el último es dependiente contrafácticamente del primero. Ambos componentes están formulados en términos de variables. El *explanans* es una proposición verdadera al efecto de que una variable del *explanandum* toma un valor particular (Cfr. Woodward, 2003, p. 202).

Es necesario precisar y dar mayor sustancia a estas ideas para entender mejor en qué consiste el modelo explicativo de Woodward. Pero, antes de pasar a la presentación más formal de su teoría hay que explicar la noción de “intervención”. De acuerdo con el propio Woodward “es heurísticamente útil pensar en una intervención como una manipulación experimental idealizada que se lleva a cabo en alguna variable X con el propósito de determinar si los cambios en X están causalmente relacionados con los cambios en otra variable Y ” (Woodward, 2003, p. 94). La idea es que una intervención (que es un tipo particular de experimentación) llevada a cabo de un modo adecuado nos permite establecer si una variable X está relacionada

causalmente con una variable Y . Presentaré un ejemplo que propone el propio Woodward para clarificar esta noción. Supongamos que realizamos un experimento para determinar si el tratamiento con un medicamento provoca que un paciente se recupere de una enfermedad. Imaginemos que tenemos una población de sujetos que padecen esa enfermedad. El experimento consiste en dividir al grupo en dos sub-grupos, uno de ellos recibirá el tratamiento con el medicamento y el otro no. Al último lo llamamos “grupo de control”. Para cada sujeto (u_i), podemos representar su estatus con respecto a si recibe o no el tratamiento con una variable T , que toma el valor 0 en caso de que no reciba el tratamiento y el valor 1 en caso de que sí lo reciba. También podemos representar su estatus de recuperación con la variable R , que toma el valor 0 si el sujeto no se recupera y el valor 1 si el sujeto se recupera. Lo que se busca saber con este experimento es si el suministro del medicamento tiene alguna relación causal con la recuperación de la enfermedad. El resultado será positivo si el grupo de sujetos que reciben el medicamento muestran una tendencia mayor a la recuperación que los sujetos que pertenecen al grupo de control. Ahora bien, la intervención en este caso consiste meramente en la asignación de los valores de T a los sujetos individuales (Cfr. Woodward, 2003, pp. 94,95).²³ En pocas palabras, la intervención consiste en la modificación de los valores de la variable X , en un contexto experimental controlado, con la finalidad de investigar si la variable X tiene algún tipo de relación causal con la variable Y . Una vez equipados con este concepto de intervención podemos entender mejor la siguiente propuesta de Woodward:

Supongamos que M es un *explanandum* que consiste en la afirmación de que alguna variable Y toma el valor particular y . Entonces un *explanans* E para M consistirá en (a) una generalización G que relacione cambios en los valores de la variable X con cambios en Y , y (b) una afirmación (de

²³ Para una explicación más detallada y precisa de en qué consiste una intervención véase el capítulo 3 de Woodward, *Making Things Happen* (2003), especialmente las páginas 104-114.

condiciones iniciales) de que la variable X toma el valor particular x . Una condición necesaria y suficiente para que E sea mínimamente explicativa con respecto a M es que (i) E y M sean verdaderas o aproximadamente verdaderas²⁴; (ii) de acuerdo con G , Y toma el valor y bajo una intervención en la que X toma el valor x ; (iii) hay una intervención que cambia el valor de X de x a x' donde $x \neq x'$, donde G describe correctamente el valor y' que Y tomaría bajo esta intervención, donde $y' \neq y$ (Woodward, 2003, p. 203).

Ilustraré esta idea con un ejemplo.²⁵ Supongamos que deseamos explicar por qué se quemó mi casa (M). Supongamos que la explicación causal abreviada de este hecho es que mi casa se quemó a causa de un corto-circuito. En este caso representamos el estatus de mi casa con la variable Y y decimos que toma el valor y en caso de que se quemó y el valor y' en caso de que no se quemó. Sea X la variable que representa al cortocircuito. Un explanans (E) exitoso de este fenómeno consiste, entonces, en: (a) la generalización que afirma que si el cortocircuito (es decir la variable X) provocara una temperatura mayor a los 200°C (es decir, si tomará el valor x), entonces la casa se incendiaría (es decir, Y tomaría el valor y); pero en caso de que el cortocircuito produjera una temperatura menor a los 200°C (es decir, si X tomara el valor x'), entonces la casa no se incendiaría (es decir, Y tomaría el valor y'). Y (b) la afirmación de que, de hecho, hubo un cortocircuito en mi casa que alcanzó una temperatura mayor a los 200°C.

Woodward presenta, entonces, una propuesta según la cual una explicación satisfactoria consiste en un conjunto de proposiciones que incluye al menos una generalización que proporciona información que muestra una dependencia contrafáctica entre los valores de las variables del *explanandum* y los valores de las variables del *explanans*. Hay que recordar que la generalización que proporciona información contrafáctica debe afirmar que los cambios en los

²⁴ Las cursivas son mías.

²⁵ A pesar de que el ejemplo que presento a continuación está demasiado simplificado debe ser suficiente para satisfacer los propósitos ilustrativos que persigo aquí: debe ser suficiente para ejemplificar cómo deben relacionarse las variables del *explanans* y el *explanandum*, mediante una generalización. Efectivamente, una explicación más sofisticada y profunda involucraría una complejidad mucho mayor a la del presente ejemplo.

valores de las variables del *explanans* están relacionados con cambios en los valores de las variables del *explanandum*. Una de las razones por la que la explicación sobre el incendio de la casa es satisfactoria consiste en que ésta proporciona la siguiente información contrafáctica, asociada con la afirmación causal “los cortocircuitos causan incendios”: “si no hubiera ocurrido un corto-circuito, la casa no se hubiera quemado”. La generalización debe afirmar que hay una relación entre algunas características de los cortocircuitos y algunas características de los incendios. Como quedó claro con el ejemplo de la casa, se debe establecer que si el cortocircuito alcanza una temperatura x , entonces habría un incendio. Es decir, la generalización afirma que se dan algunas relaciones entre algunas características del *explanans* (el cortocircuito) y algunas características del *explanandum* (el incendio). Woodward insiste en que para que una explicación sea exitosa no basta con que la generalización afirme que tales relaciones se dan, sino que además la generalización “debe describir correctamente cómo la variable del *explanandum* cambiaría bajo intervenciones en las variables del *explanans*” (Woodward, 2003, p. 202). Es decir, las relaciones de cambio entre los valores de las variables del *explanans* y los del *explanandum* que se presentan en la generalización deben ser correctas, deben representar *verídicamente* el modo en que de hecho se relacionan los valores de estas variables.

En este punto se puede ver claramente cómo el éxito de una explicación causal depende de la verdad de ciertas proposiciones. Por lo tanto, el tipo de entendimiento científico que surge de una explicación causal adecuada depende, a su vez, de ciertas creencias verdaderas. En particular depende de que sea verdadera la creencia en la generalización presente en toda explicación causal. De este modo, el valor del tipo de entendimiento científico que surge de la explicación causal es explicable también en términos de la TVEE, pues ejemplifica la relación de dependencia con la creencia verdadera, que sostiene que una unidad de evaluación ejemplifica la relación de *dependencia* con la creencia verdadera cuando su valor depende de

que otras creencias (u otras actitudes dóxicas) sean verdaderas o bien tengan un valor epistémico distinto de la verdad que sea explicable en términos de la TVEE. Al igual que con los tipos de explicación que se analizaron anteriormente, si el entendimiento que surge de explicaciones causales no ejemplificara la relación de dependencia con la creencia verdadera no lo consideraríamos epistémicamente valioso.

Con el análisis de esta teoría contrafáctica de la explicación causal concluyo con los análisis dedicados a los tipos de entendimiento científico que surgen de explicaciones científicas. He mostrado que los tres tipos de entendimiento ejemplifican al menos una de las relaciones adecuadas con la verdad que presenté en la sección (1.2) y, además, que si no ejemplificaran esas relaciones entonces no los consideraríamos epistémicamente valiosos. A continuación analizaré un tipo de entendimiento que surge de la unificación teórica.

2.2 La unificación teórica

Antes de analizar con detalle la forma de entendimiento que surge de la unificación teórica me gustaría presentar un ejemplo histórico de la ciencia en el que se alcanzó entendimiento científico de un fenómeno particular mediante unificación teórica.

El descubrimiento de la fisión nuclear

Enrico Fermi, un físico de origen italiano considerado junto con J. Robert Oppenheimer el padre de la bomba atómica, realizó en 1934 experimentos en los que bombardeaba muestras de uranio con neutrones. En aquel momento se creía que los elementos con número atómico mayor

a noventa y dos (también conocidos como elementos transuránicos y que se caracterizan por ser inestables y radiactivos) se formaban al bombardear átomos de uranio con neutrones.

Sin embargo, en 1938, cuando Otto Hahn y Fritz Straßmann, mientras replicaban el experimento de Fermi bombardeando átomos de uranio con neutrones lentos obtenidos de una mezcla de radio y berilio, observaron algunos resultados inesperados (*Cfr.*, Lapp, 1969, p. 162). En esa época aún no se había establecido la existencia de los actínidos y se suponía erróneamente que el uranio (que de hecho es un miembro de la familia de los actínidos) era un elemento del grupo seis. Por lo tanto, se creía que los primeros elementos transuránicos debían ser similares a los elementos de los grupos siete a diez, es decir se creía que serían similares al renio y a los elementos del grupo del platino (esto es, el rutenio, el rodio, el paladio, el osmio, el iridio y el platino). De hecho, Hahn y Straßmann establecieron en sus experimentos la presencia de cuatro isótopos de estos elementos, pero los identificaron erróneamente como los elementos 93-96. Lo más relevante, sin embargo, es que también encontraron evidencia de la presencia de algún isótopo de algún metal alcalinotérreo (del grupo dos). Este descubrimiento representaba un problema pues no encajaba con los otros elementos que aparecían.

En un inicio, Hahn sospechaba que habían encontrado radio y que éste se había formado al separar dos partículas alfa del núcleo del uranio. En ese tiempo, el consenso científico decía que lograr esto mediante el bombardeo con neutrones era poco probable. El mismo Niels Bohr señaló esta cuestión a Hahn y sugirió que quizás se trataba de un elemento transuránico extraño (*Cfr.* Rife, 2007, pp. 180-181). Pero, por otro lado, si no se trataba de uranio, entonces debía tratarse de otro metal alcalinotérreo, por ejemplo bario. Pero esta alternativa resultaba mucho peor: creer que los isótopos eran de bario sugería que debía existir un proceso mediante el cual se transforma el uranio en bario, lo cual implicaba remover cien nucleones (protones o neutrones) del núcleo del uranio, algo que simplemente resultaba absurdo para la física aceptada en esa época: “según las teorías aceptadas por la física nuclear, eso era algo tan

inimaginable como sería romper un ladrillo disparándole un guisante” (Lapp, 1969, p. 162). Algunos refinamientos en la técnica experimental llevaron a confirmar la hipótesis más desafortunada: que los isótopos de elementos alcalinotérreos se comportaban consistentemente como bario, no como radio. El propio Hahn escribió una carta a Lise Meitner, una antigua colaboradora que jugó un papel muy importante en esta historia, confirmando sus sospechas: “cada vez nos acercamos a la conclusión atroz de que nuestros isótopos de Ra [radio] no se comportan como Ra, sino como Ba [bario]” (Rife, 2007, p. 183). Hasta este punto, los descubrimientos de Hahn y Straßmann resultaban simplemente enigmáticos e inexplicables: la teoría aceptada en aquella época no tenía modo de dar cuenta de semejante fenómeno. Constituía un auténtico fenómeno sorprendente y no había un modo de incorporarlo a la teoría atómica establecida. Hahn y Straßmann, siendo químicos, evitaron dar alguna explicación física del fenómeno observado, si bien Hahn llegó a especular que la fisión, el proceso nuclear por el que el núcleo de un átomo se descompone en partes más pequeñas, era la única explicación razonable; sin embargo era incapaz de dar una explicación física completa de cómo sucedía esto²⁶.

En diciembre de 1938 Meitner, quien estaba al tanto de los resultados de los experimentos de Hahn y Straßmann, y su sobrino Otto Frisch fueron capaces de dar una elucidación matemática del proceso de fisión nuclear haciendo uso del modelo del núcleo atómico como gota líquida, ideado por primera vez por Bohr.²⁷ Es cierto que Meitner fue principalmente

²⁶ Aquí se presenta la historia de un modo simplificado. En realidad los implicados en el desarrollo de esta teoría tuvieron serias dudas a la hora de postular una teoría sobre el núcleo atómico que desafiara las teorías más establecidas al respecto en ese momento. Especialmente Hahn y Leiter tuvieron serios problemas al apoyar esta teoría, pues implicaba aceptar que el trabajo que venían haciendo desde hacía varios años sobre los elementos transuránicos tendría que ser reformulado y, quizás, rechazado completamente. Para una explicación detallada del proceso intelectual que llevó a proponer la teoría de la fisión nuclear pueden consultarse los capítulos 9 y 10 de Rife (2007).

²⁷ De hecho el propio Frisch tuvo la oportunidad de platicar con Niels Bohr en enero de 1939, una vez que Meitner y Frisch llevaban tiempo desarrollando su explicación de la fisión nuclear (aunque aún no la publicaban). Frisch le comunicó qué teoría estaban desarrollando y le explicó que haría uso de su modelo del núcleo atómico de la gota de agua. Según una carta de Frisch a Meitner “la conversación duro sólo cinco minutos, pues Bohr inmediatamente y en todos los aspectos estuvo en acuerdo con nosotros. Sólo estaba sorprendido porque él no

responsable por el desarrollo del modelo matemático que explicaba la cuestión, pero Frisch, quien sí tenía acceso a equipo de laboratorio sofisticado, llevó a cabo experimentos que resultaron cruciales para probar la teoría que recién habían desarrollado sobre la fisión nuclear.

De acuerdo con la teoría física establecida, el núcleo atómico se mantiene unido por la así llamada “interacción nuclear fuerte”, una de las cuatro interacciones fundamentales que sirven para explicar la interacción entre partículas en el modelo estándar de la física de partículas. Esta fuerza es precisamente la responsable de mantener unidos a los neutrones y protones en el núcleo atómico. Como los protones tienen carga positiva, esta fuerza debe superar la repulsión eléctrica entre estas partículas. Sin embargo, la interacción nuclear fuerte actúa sólo a una corta distancia: una fracción del diámetro de los núcleos atómicos más grandes. Es por esto que no existen naturalmente de manera estable elementos con números atómicos mayores a los del uranio: pues cuando el núcleo atómico es mayor al del uranio, las fuerzas eléctricas repelentes, que actúan a mayores distancias, predominan sobre la fuerza nuclear. Ahora bien, de acuerdo con la explicación de la fisión nuclear que propusieron Meitner y Frisch, que se basa en el modelo del núcleo de la gota líquida, esto es lo que sucede en el proceso de fisión nuclear:

Cuando un neutrón externo penetra el “líquido” del núcleo, la perturbación hace oscilar violentamente al núcleo. Al llegar a cierto punto se deformará adquiriendo una forma alargada, y cuando esto ocurra, las fuerzas cohesivas de corto alcance del núcleo ya no serán lo suficientemente poderosas para mantenerlo unido. Las fuerzas eléctricas repelentes dominan y comienzan a empujar los dos extremos del núcleo, apartándolos cada vez más. Formará como una cintura de avispa, y acabará por dividirse en dos mitades, de modo muy parecido a como se dividen las células biológicas (Lapp, 1969, pp. 63-64).

había considerado antes esta posibilidad, que se sigue tan directamente de las concepciones presentes de la estructura nuclear. También está completamente de acuerdo con nuestra postura de que la desintegración de un núcleo pesado en dos piezas es casi un proceso clásico, que no ocurre en absoluto debajo de cierta energía, pero que ocurre muy fácilmente un poco sobre ella” (Rife, 2007, p.197).

La teoría propuesta por Meitner y Frisch no sólo daba cuenta del fenómeno observado en los experimentos conducidos por Hahn y Straßmann, sino que incluso logró llevar a cabo predicciones exitosas que le dieron un mayor sustento empírico. Si efectivamente lo que había sucedido en los experimentos en cuestión había sido un caso de fisión nuclear, entonces la teoría de Meitner y Frisch predecía que también debía observarse otro fenómeno: una gran liberación de energía. De acuerdo con los cálculos hechos por Meitner y Frisch, la energía liberada por cada átomo de uranio dividido sería de doscientos millones de electrón voltios (*Cfr.* Rife, 2007, p. 190). Lo cual efectivamente ocurrió en los experimentos que se condujeron posteriormente para probar esta predicción: “[Frisch] bombardeó el uranio con neutrones. A medida que cada neutrón se estrellaba en un átomo de uranio, observaba una gran explosión de energía, lo bastante potente para hacer salir de escala a los instrumentos de medida” (Lapp, 1969, p. 164).

Así, este caso presenta paradigmáticamente una situación en la que se debe modificar una teoría aceptada para poder incorporar un fenómeno sorprendente. Hahn y Straßmann alcanzaron unos resultados experimentales que no eran explicables de acuerdo con la teoría física contemporánea aceptada en su época. Se enfrentaban a un fenómeno misterioso para el cual no podían dar ninguna explicación. En otras palabras, no entendían el fenómeno: sabían que bombardear el uranio con neutrones resultaba en algunos átomos de bario, pero no entendían por qué sucedía esto. Hizo falta que Meitner y Frisch dieran una teoría física de la naturaleza del núcleo atómico que daba una explicación del fenómeno en cuestión que implicaba desafiar algunos supuestos fundamentales de la teoría física contemporánea. Además hizo falta que se acumulara evidencia empírica a su favor. Una vez que se desarrolló la teoría de la fisión atómica de Meitner y Frisch, entonces se estuvo en posición de entender científicamente el fenómeno observado en los experimentos de Hahn y Straßmann. Es muy importante señalar que

no todo caso de entendimiento científico implica el desarrollo de nuevas teorías y el desafío radical de viejas teorías. Presento este caso simplemente para ilustrar un modo particular de alcanzar entendimiento científico. Más adelante, en la discusión de la teoría de Schurz y Lambert este ejemplo será bastante útil para ilustrar cómo funcionaría su teoría.

El modelo de Schurz y Lambert

Hasta aquí he analizado tres tipos de entendimiento científico que surgen de distintas teorías de la explicación científica y he explicado cómo el valor epistémico de esas formas de entendimiento puede explicarse desde un punto de vista veritista. Además, al analizar cada caso particular, señalé cómo se ha relacionado tradicionalmente el concepto de explicación científica con el de entendimiento científico: quienes proponen los distintos modelos explicativos han insistido en que tener una explicación científica satisfactoria de un fenómeno nos permite entenderlo. En lo que sigue, abordaré una concepción del entendimiento que, a diferencia de las anteriores, pone el énfasis en la noción de *unificación teórica*. Centraré mi atención en la propuesta que presentan Gerhard Schurz y Karel Lambert (1994) y, del mismo modo que hice en los casos anteriores, investigaré si es posible explicar el valor epistémico de esta forma de entendimiento científico en términos de la TVEE. Las teorías que he revisado hasta aquí adoptan la tesis de que las explicaciones son el medio para alcanzar el entendimiento. En la teoría que defienden Schurz y Lambert, el medio por el que se alcanza el entendimiento es la unificación teórica que se alcanza mediante la incorporación de un fenómeno a un *corpus* cognitivo. En otras palabras, Schurz Y Lambert sostienen la idea de que es posible alcanzar entendimiento científico sin apelar a explicaciones científicas.²⁸

²⁸ Peter Lipton, por ejemplo, argumenta enérgicamente a favor de la idea de que la explicación científica no es el único medio que nos puede proporcionar entendimiento científico. Por ejemplo, las ilustraciones, los modelos y las intervenciones pueden darnos genuino entendimiento científico, desde el punto de vista de Lipton. Al inicio del siguiente capítulo diré algo sobre su postura. Si se desea profundizar sobre este punto, véase su

Schurz y Lambert ofrecen una caracterización del entendimiento científico que se basa en la noción de incorporación de un fenómeno a un corpus cognitivo. De acuerdo con ellos, entender un fenómeno F no es otra cosa sino saber de qué modo F se incorpora en el conocimiento previo que uno tiene (Cfr. Schurz y Lambert, 1994, p. 66). De modo más preciso, el entendimiento es una relación que se da entre tres elementos: una respuesta R a una pregunta que busca entendimiento, un fenómeno F y un corpus cognitivo C . Se dice que R da entendimiento de F cuando R incorpora *de manera exitosa* a F en C . Es importante señalar que la incorporación se lleva a cabo siempre mediante el uso de un “argumento en sentido amplio”,²⁹ de modo que decir que F se incorpora en C siempre es una expresión abreviada de “ F se incorpora a C mediante una inferencia I ”. Quizás sea importante aclarar qué quieren decir los autores con la noción de *corpus cognitivo*. Según ellos, un corpus cognitivo está constituido por “todas las oraciones conocidas o creídas por el investigador (incluyendo las oraciones de observación, leyes, teorías e hipótesis con distinto grado de credibilidad)” (Schurz y Lambert, 1994, p. 67). Los *corpora* cognitivos pueden a su vez dividirse en un subconjunto K , que incluye a todas las oraciones que describen los fenómenos que el sujeto conoce y en un subconjunto I , que incluye a todas las inferencias que el sujeto domina.

Una ilustración ayudará a clarificar esta propuesta. Supongamos que buscamos incorporar a nuestro corpus cognitivo el fenómeno de que el agua de mar no es útil para el consumo humano. La razón por la que los seres humanos no pueden beber agua de mar consiste en que la concentración de sal en el agua de mar es superior a lo que pueden procesar los riñones. Es decir, para eliminar la sal ingerida al beber agua de mar los riñones deberían producir un volumen mayor al del agua que se bebe, de modo que beber agua de mar no ayuda a la hidratación. De este modo el fenómeno se incorpora al corpus cognitivo mediante la siguiente

“Understanding without Explanation” (2009).

²⁹ Más adelante explicaré con detalle qué es un argumento en sentido amplio.

inferencia: 1) sólo el agua con una concentración de sal menor a x es útil para el consumo humano; 2) el agua de mar tiene una concentración de sal mayor a x ; por lo tanto, 3) el agua de mar no es útil para el consumo humano.

Sin embargo, como bien señalan los autores, la manera de hablar de incorporación de un fenómeno a un corpus cognitivo puede resultar un poco engañosa, ya que al introducir un nuevo elemento F en C , el corpus cognitivo mismo cambia. Por lo tanto hay que distinguir entre el corpus cognitivo previo a la incorporación de F , llamémoslo C , y el corpus cognitivo posterior a la incorporación de F , llamémoslo C^* (que es igual a $C+F$). Esta precisión resulta bastante importante para el modelo que quieren presentar los autores, pues en última instancia la manera en que se determina si un fenómeno ha sido exitosamente incorporado, y si su incorporación contribuye entendimiento, consiste en comparar el grado de unificación entre C y C^* . Así, cuando decimos que F se incorpora exitosamente a C , queremos decir que la introducción de F en el corpus cognitivo incrementa la unificación de C . En otras palabras, no basta con que el fenómeno incorporado sea consistente con el corpus cognitivo, para que haya una incorporación genuina, además hace falta que la incorporación de F tenga un impacto positivo en la unificación de C .

A continuación presentaré de un modo más formal y preciso la propuesta de Schurz y Lambert. Esto permitirá dar mayor sustancia y precisión a la noción de incorporación y de unificación que son tan importantes en su teoría. Pero antes hay que explicar qué es un “argumento en sentido amplio”, pues esta noción tiene un papel importante en la precisión formal de la noción de incorporación. De acuerdo con los propios autores:

Un argumento en sentido amplio es cualquier par formado por un conjunto de oraciones *Prem* (las premisas) y una oración *Con* (la conclusión) donde una cantidad “suficiente” de información se “transmite” de *Prem* a *Con*.

Esto se denota con $Prem \rightarrow Con$. En términos probabilistas, esto significa que la probabilidad de Con se incrementa a causa de $Prem$ a un valor “suficientemente” grande (Schurz y Lambert, 1994, p. 71).

Así, la característica central de todo argumento en sentido amplio es que la probabilidad de la conclusión se incrementa, dada la información de las premisas. Podemos expresar esta relación del siguiente modo:

$$P(Con) < P(Con|Prem)$$

Más adelante, cuando se explore el modo en que la teoría veritista podría explicar esta forma de entendimiento en sus términos, analizaré con mayor detalle la noción de argumento en sentido amplio. Por el momento basta con señalar que, entre los argumentos en sentido amplio se destacan los argumentos deductivos e inductivos. Ya que he explicado, en términos generales, qué es un argumento en sentido amplio, podemos entender mejor la noción de incorporación que presentan los autores:

Un argumento en sentido amplio (que tiene la forma $Prem \rightarrow Con$) incorpora a un fenómeno F de manera exitosa en C^* si y sólo si $Prem \rightarrow Con$ es un conector en C^* , que contiene a F como la conclusión (esto es $F=Con$) o en el conjunto de premisas (esto es, $F \in Prem$) tal que C^* está más unificado que C (Cfr. Schurz y Lambert, 1994, p. 77).

Es decir, un argumento incorpora exitosamente a un fenómeno sólo si la incorporación del fenómeno aumenta la unificación del corpus y, además, el argumento es un conector en C^* y F está en las premisas o es la conclusión.³⁰ Pero para comprender completamente esta definición hace falta que a su vez se explique qué es la unificación de un corpus cognitivo y qué significa

³⁰ En última instancia, la noción de incorporación será útil para explicar la noción de entendimiento. Según los autores, el tipo de entendimiento que nos otorga la incorporación de F puede ser de dos tipos, dependiendo del lugar que F ocupe en el argumento que lo incorpora a C^* : decimos que hay *entendimiento sobre F* cuando F forma parte de las premisas del argumento que incorpora a F ; decimos que hay *entendimiento de por qué F* cuando F es la conclusión del argumento (Cfr. Schurz y Lambert, 1994, p. 73).

que un argumento en sentido amplio sea un conector en un corpus cognitivo particular. Dejaré para más adelante la explicación de la noción de unificación. Por el momento explicaré qué es un conector. De acuerdo con los propios autores:

Prem → *Con* es un *conector* en *C* si y sólo si (i) $(\text{Prem} \rightarrow \text{Con}) \in I$, (ii) $\text{Prem} \subseteq K$, $\text{Con} \in K$ y $\text{Con} \notin \text{Prem}$, (iii) no hay un subconjunto propio *Prem** de *Prem* tal que $\text{Prem}^* \rightarrow \text{Con}$ satisfaga las cláusulas (i) y (ii). (Schurtz y Lambert, 1994, p. 77).

Es decir, un argumento es un conector en un corpus particular si y sólo si: (i) la forma argumental pertenece a la clase de inferencias que domina el sujeto, es decir, al subconjunto *I* del corpus; (ii) tanto las premisas como la conclusión están en el subconjunto *K* del corpus y la conclusión no es un elemento de las premisas; y (iii) no hay un subconjunto de las premisas (llamémoslo *Prem**) que pueda reemplazar a las premisas originales del argumento y que cumpla las cláusulas (i) y (ii). Es importante señalar explícitamente cuál es la importancia de las cláusulas (ii) y (iii). (ii) tiene el objetivo de impedir que se den argumentos circulares con la supuesta finalidad de incrementar la unificación del corpus. Como veremos más adelante, uno de los requisitos que debe cumplir un corpus cognitivo para estar unificado consiste en que no haya argumentos circulares en su interior. La inclusión de esta cláusula garantiza que la unificación de un corpus cognitivo nunca se incrementará apelando a argumentos circulares, como “{T, P & Q} → P”. Por otro lado, (iii) tiene el objetivo de garantizar que todas las premisas que forman parte de un argumento en sentido amplio sean necesarias para inferir la conclusión: si hay un subconjunto de *Prem* (a saber, *Prem**) que satisface los criterios (i) y (ii), entonces es posible llevar a cabo la inferencia prescindiendo de alguno de los elementos de *Prem*. En cambio, si se cumple (iii), todos los elementos de *Prem* serán necesarios para llevar a cabo la inferencia.

Una vez que se han dado estas nociones, es posible explicar de un modo más preciso qué significa entender un fenómeno:

Una aseveración R aporta entendimiento de F a C si y sólo si R es una respuesta adecuada a la pregunta “¿Cómo se incorpora a F en C^* ?, donde el fenómeno F está incluido en el corpus cognitivo C (Schurz y Lambert, 1994, p. 75).

Es decir, uno entiende un fenómeno si tiene una respuesta adecuada a la pregunta sobre cómo se incorpora F en C^* . Ahora bien, recordemos que más arriba se dijo que la incorporación de un fenómeno se lleva a cabo siempre mediante un argumento en sentido amplio. Por lo tanto, una respuesta adecuada a la pregunta por la incorporación de F debe incluir la especificación de que existe un argumento en sentido amplio que incorpora a F en C^* :

R es una respuesta adecuada a la pregunta “¿Cómo se incorpora a F en C^* ?” si y sólo si R incluye la afirmación de que un argumento en sentido amplio es sólido y este argumento incorpora a F en C^* (Schurz y Lambert, 1994, p. 76).

Ya expliqué antes qué significa que un argumento incorpore a un fenómeno en un corpus cognitivo. Así tenemos una definición completa y precisa de qué significa entender un fenómeno. Resumiendo: entendemos un fenómeno cuando tenemos una respuesta adecuada a cómo ese fenómeno se incorpora a un corpus cognitivo; esta respuesta debe consistir en un argumento en sentido amplio correcto que incorpora exitosamente el fenómeno en el corpus. Esto último sólo se logra en última instancia si el argumento es un conector y la incorporación del fenómeno tiene como resultado un incremento de la unificación del corpus cognitivo. Hasta aquí ya se han explicado todas las nociones involucradas con la definición de entendimiento

científico, con excepción de la noción de unificación. A continuación explicaré con mayor detalle en qué consiste ésta.

Un análisis detallado muestra que, de acuerdo con los autores, la noción de coherencia tiene un papel sobresaliente en la definición de la noción de unificación. Schurz y Lambert no dan una explicación precisa de cómo entienden el concepto de coherencia, sólo señalan que la coherencia de un corpus cognitivo tiene que ver con su grado de “conectabilidad”, es decir con las conexiones que hay entre los elementos. Esta elucidación no es informativa en absoluto. El error de hacer uso de una noción vaga de coherencia se comete a menudo. Paul Thagard, uno de los filósofos que se ha esforzado por presentar una definición detallada del concepto de coherencia sostiene que, a pesar de la popularidad de las teorías coherentistas del conocimiento, la naturaleza misma de la coherencia usualmente resulta vaga, lo cual nos deja sin un método preciso para determinar cuándo una creencia contribuye coherencia a un sistema o, al contrario, la disminuye (*Cfr.* Thagard, *et al.*, 2002, p. 104). Dado que Schurz y Lambert no señalan explícitamente qué noción de coherencia utilizan en su teoría, me tomaré la libertad de suponer que asumen la noción que propone el propio Thagard. Los elementos pueden ser coherentes (adecuarse) o ser incoherentes (resistir a la adecuación) dependiendo de qué tipo de relación se da entre ellos. Las relaciones de coherencia incluyen explicación, deducción, similitud, asociación, etc. Las relaciones de incoherencia incluyen inconsistencia, incompatibilidad y asociación negativa. Así, entre mayores relaciones de coherencia (explicación, deducción, similitud, asociación) se den entre los elementos y entre menos relaciones de incoherencia (inconsistencia, incompatibilidad y asociación negativa) se den, entonces el sistema será más coherente.³¹

³¹ Para una explicación detallada de la teoría de la coherencia de Thagard véase Thagard, *et al.*, “Knowledge and Coherence” (2002).

Sin embargo, a pesar de que la coherencia tiene un papel sobresaliente en la unificación, no lo es todo. Los autores descartan dar una elucidación de la unificación en términos de mera coherencia. Por esta razón, favorecen una elucidación según la cual la unificación es coherencia más la satisfacción de tres requisitos que surgen de consideraciones sobre la actividad científica. Esto se debe a que su finalidad es dar una elucidación del entendimiento *científico* y, como tal, deben tomarse en cuenta algunas particularidades de la actividad científica. En particular, insisten en que una elucidación adecuada de la noción de unificación (en el contexto científico) debe garantizar que los *corpora* cognitivos no sólo sean coherentes, sino que además estén *confirmados empíricamente*. El hecho de que la confirmación empírica esté asegurada en un corpus cognitivo bien formado garantiza que no habrá casos en los que la unificación de un corpus se incrementa mediante la *especulación sin contenido empírico* (Cfr. Schurz y Lambert, 1994, p. 85). Si se busca incrementar la coherencia de un corpus, pero se hace mediante especulaciones que no tienen ningún tipo de confirmación empírica, entonces no se incrementará la unificación del corpus, a pesar de que sí incrementa su coherencia. Las primeras dos condiciones adicionales se introducen precisamente para garantizar la confirmación empírica de los *corpora* cognitivos³², de acuerdo con los propios autores “las dos restricciones garantizan que la unificación científica arroja como subproducto confirmación empírica” (Schurz y Lambert, 1994, p. 72):

1. “Que los fenómenos estén conectados de un modo *no-circular* vía argumentos en sentido amplio. Efectivamente, si éste es el caso, entonces muchos fenómenos están conectados con, y por lo tanto son al menos parcialmente reducibles a, un conjunto pequeño de fenómenos básicos no conectados.

³² La tercera condición, como veremos más adelante, tiene que ver más bien con un principio metodológico general de la ciencia.

2. “[...] que los fenómenos estén conectados de un modo no circular vía argumentos en sentido amplio siempre tiene que contribuir a la unificación del subconjunto de los fenómenos llamados los *datos*” (Schurz y Lambert, 1994, p. 72).

Los autores introducen una tercera condición de manera no explícita. Hay casos en los que a pesar de que una respuesta R aumente la coherencia del corpus cognitivo y cumpla las condiciones (1) y (2), ésta no proporciona entendimiento científico. Se refieren a los así llamados hechos desconcertantes, es decir hechos que todo intento imaginable de incorporación (a pesar de satisfacer (1) y (2)) involucra una respuesta aún más desconcertante que el hecho inicial que se deseaba incorporar. Por ejemplo, supongamos que estamos en un mercado y vemos pasar un elefante. Este hecho es desconcertante. Y cuando preguntamos qué hace ahí un elefante alguien nos explica que el mercado fue designado hace algunos meses un parque de animales. Esta respuesta no proporciona genuino entendimiento porque, a pesar de incrementar la coherencia del corpus y satisfacer (1) y (2), simplemente se intercambia un hecho desconcertante por otro. Este caso, además, satisface las cláusulas (1) y (2): la respuesta que se da para incorporar el hecho desconcertante al corpus no es circular y, además, contribuye a la unificación del conjunto de los datos, pues la presencia de un elefante en el mercado está registrado en el conjunto de los datos, es decir es un fenómeno que se conoce mediante observación.

Para evitar que la incorporación de hechos desconcertantes cuente como una que incrementa la unificación del corpus y que proporciona entendimiento genuino, los autores introducen una tercera condición que debe cumplir una incorporación genuina de un fenómeno para incrementar la unificación:

3. La incorporación de un fenómeno desconcertante a un corpus cognitivo no puede llevarse a cabo mediante el uso de un argumento en sentido amplio que haga uso, a su vez, de un fenómeno igualmente desconcertante.

Para que esta condición sea informativa hace falta decir con precisión cuándo un fenómeno es *desconcertante*. Según Schurtz y Lambert, decir que un fenómeno es desconcertante no significa meramente que es *improbable*. Un fenómeno desconcertante es uno que “está en conflicto con K^{33} y, además, toda 'causa' imaginable o 'razón' para explicarlo, de acuerdo con las leyes y teorías de K , entra en conflicto con K ” (Schurtz y Lambert, 1994, p. 96). Sin embargo, para que esta definición esté completa hace falta especificar qué significa que un fenómeno esté “en conflicto” con K : “un fenómeno singular está en *conflicto con K* si y sólo si es *improbable*, dado el resto de K [...]. Un fenómeno singular F en K es *improbable*, dado el resto de K , si y sólo si existe un argumento inductivo en I que tiene a $\neg F$ como conclusión” (Schurtz y Lambert, 1994, p. 96). Es decir, un fenómeno F está en conflicto con K cuando es improbable, desde el punto de vista de K , en el sentido de que hay un argumento inductivo que utiliza sólo premisas de K y que tiene como conclusión $\neg F$. Finalmente, decimos que un hecho es desconcertante cuando él mismo está en conflicto con K y cualquier modo imaginable de explicarlo apela, a su vez, a un hecho que también está en conflicto con K .

Para ilustrar la noción de *hecho desconcertante* echaré mano del ejemplo que presenté al inicio de esta sección sobre el descubrimiento de la fisión nuclear. Recordemos que Hahn y Straßmann, intentando descubrir elementos transuránicos, llevaron a cabo experimentos en los que bombardeaban muestras de uranio con neutrones. En estos experimentos observaron un resultado inesperado: la presencia de algún isótopo de algún metal alcalinotérreo.

³³ K es el subconjunto del corpus cognitivo que contiene todas las oraciones que describen los fenómenos que el sujeto conoce.

Posteriormente lograron identificar a las moléculas como isótopos de bario. Este hecho, es decir que el bombardeo de uranio con neutrones tuviera como resultado la aparición de isótopos de bario, resultaba precisamente desconcertante: de acuerdo con las teorías físicas aceptadas en la época semejante hecho era simplemente inexplicable. La física de la época no sólo no tenía ningún modo de dar cuenta de ese hecho, sino que incluso el hecho mismo era altamente improbable a los ojos de las teorías aceptadas: “según las teorías aceptadas por la física nuclear, eso era algo tan inimaginable como sería romper un ladrillo disparándole un guisante” (Lapp, 1969, p. 162). De manera que el hecho de que aparecieran isótopos de bario no sólo era improbable, dada la teoría física del momento, sino que cualquier intento de explicar el hecho apelaba, a su vez, a algún hecho desconcertante. Este conflicto sólo fue superado una vez que Lise Meitner desarrolló un nuevo modelo matemático de la estructura nuclear que permitía explicar cómo los átomos (especialmente los transuránicos) podían dividirse mediante el bombardeo de neutrones. Sin embargo, en ausencia de tal innovación teórica, los físicos de la época se encontraban con un hecho genuinamente desconcertante.

Así, para que la incorporación de un fenómeno incremente la unificación del corpus y que, por lo tanto, proporcione genuino entendimiento no basta con el mero incremento de coherencia, ni tampoco basta con la coherencia más la satisfacción de las condiciones (1) y (2). Además, debe cumplirse la condición (3). Cuando la incorporación de un fenómeno, además de aumentar la coherencia del corpus cognitivo, cumple las tres condiciones señaladas, decimos que también incrementa la unificación del corpus. En última instancia, esta característica (el incremento de la unificación) es la que determina si un fenómeno se entiende genuinamente.

Resumiendo, Schurz y Lambert proponen lo siguiente: uno entiende un fenómeno F si y sólo si tiene a la mano una respuesta adecuada R para la pregunta “¿Cómo se incorpora F en el corpus cognitivo C^* ?” La respuesta R debe consistir en un argumento en sentido amplio correcto que incluye a F como premisa o conclusión, incorpora exitosamente al fenómeno en el

corpus y que tiene como resultado que C^* esté más unificado que C . Ahora bien, la unificación de un corpus cognitivo es igual a su coherencia, siempre y cuando se satisfagan estos requisitos adicionales: 1) no circularidad, 2) confirmación por los datos y 3) no apelar a fenómenos desconcertantes. Omitiendo muchos detalles podríamos sintetizar su proposición diciendo que uno entiende un fenómeno cuando su incorporación (mediante un argumento en sentido amplio) incrementa la unificación del corpus cognitivo.

En este punto uno podría preguntarse cómo se determina de un modo exacto cuándo un corpus cognitivo está más unificado que otro. Schurz y Lambert presentan un sofisticado método para elaborar un análisis comparativo del grado de unificación entre cualesquiera dos *corpora* cognitivos. Entre otras características, el método asigna *pesos* a los fenómenos dependiendo de si son básicos o derivados y premia con un grado mayor de unificación a los *corpora* con un menor número de fenómenos básicos y mayor número de fenómenos derivados incorporados. Dado mis propósitos no es necesario profundizar sobre el modo en que se determina cuál de dos corpus cognitivos está más unificado. Basta con entender en términos generales cuál es la propuesta que presentan Schurz y Lambert.³⁴

Aquí asumiré que efectivamente se puede establecer (utilizando el método que presentan los autores) si la incorporación de un fenómeno aumenta o disminuye la unificación del corpus cognitivo. De acuerdo con los autores, uno entiende F sólo si su incorporación incrementa la unificación del corpus: “Nuestro tratamiento no requiere que una respuesta R a una pregunta que busca entendimiento produzca unificación global, sino simplemente que incremente la unificación” (Schurz y Lambert, 1994, p. 74).

La pregunta que se debe responder ahora es si esta noción de entendimiento científico es también explicable según los términos veritistas. Los propios autores le dan un papel relevante

³⁴ Para una presentación detallada del modo en que se calcula el grado de unificación de un corpus, con respecto a otro, véase la sección 2.2 (pp. 78-85) y el apéndice de “Outline of a Theory of Scientific Understanding” (1994), de Schurz y Lambert.

en su propuesta a los argumentos en sentido amplio. Se recordará que los argumentos en sentido amplio entran en la propuesta del siguiente modo: toda incorporación exitosa de un fenómeno F en un corpus cognitivo se lleva a cabo *mediante* un argumento en sentido amplio, que incluye a F , ya sea como una de las premisas o como la conclusión. Pero los autores rehúsan dar condiciones de suficiencia para la corrección de los argumentos en sentido amplio y en su lugar dan una lista de tipos de argumentos en sentido amplio correctos. La lista se encuentra constituida por los argumentos deductivos, los argumentos inductivos y los argumentos aproximativos.³⁵ Ahora bien, a pesar de no dar una condición de corrección para todos estos tipos de argumentos, los autores sí presentan una caracterización muy general de los argumentos en sentido amplio:

Un argumento en sentido amplio es cualquier par formado por un conjunto de oraciones *Prem* (las premisas) y una oración *Con* (la conclusión) donde una cantidad “suficiente” de información se “transmite” de *Prem* a *Con*. Esto se denota con $Prem \rightarrow Con$. En términos probabilistas, esto significa que la probabilidad (grado de creencia) de *Con* se incrementa a causa de *Prem* a un valor “suficientemente” grande (Schurz y Lambert, 1994, p. 71).

Así, tenemos que todo argumento en sentido amplio correcto es un argumento tal que la probabilidad de la conclusión se incrementa sustancialmente a la luz de la información proporcionada por las premisas. De acuerdo con esta teoría entender un fenómeno F consiste en tener un argumento en sentido amplio en el que F está en las premisas o la conclusión y que incrementa a un *grado alto* la probabilidad de la conclusión. Dada la naturaleza de los

³⁵ De acuerdo con los autores “las inferencias aproximativas son aplicaciones indirectas de inferencias deductivas válidas: $Prem \rightarrow_a Con$ es correcta si y sólo si hay una Con^* tal que $Prem$ implica lógicamente Con^* y Con^* se aproxima a Con ” (Schurz y Lambert, 1994, p. 91). Para entender completamente en qué consiste una inferencia aproximativa hace falta explicar la noción de que una proposición se aproxime a otra. Sin embargo, para los propósitos de esta investigación basta con señalar que este tipo de inferencias hacen uso de inferencias deductivas, por lo tanto, bastará con analizar qué sucede con las inferencias inductivas y deductivas.

argumentos en sentido amplio, se sigue que la actividad de entender un fenómeno mediante la incorporación de éste a un corpus cognitivo ejemplifica la relación de ser altamente probable, dada la evidencia. Pues si uno razona del modo que prescriben los argumentos en sentido amplio, entonces adquiere creencias que son altamente probables, dada la evidencia que se presenta en las premisas. Así, el entendimiento que surge de este modo ejemplifica la relación de ser altamente probable dada la evidencia que se presentó en la sección (1.2).

Sin embargo, hay que recordar que si bien la incorporación de un fenómeno a un corpus es una condición necesaria para entender un fenómeno, no es una condición suficiente. Para que haya genuino entendimiento, la incorporación debe incrementar la unificación del corpus. En particular, debe de incrementar su coherencia y satisfacer los requisitos (1)-(3). Por lo tanto si queremos explicar el valor epistémico de esta forma de entendimiento científico no basta con que expliquemos el valor de la incorporación de un fenómeno en términos de la TVEE. Hace falta que también expliquemos de qué modo el valor epistémico de los requisitos (1)-(3) puede explicarse en términos de la teoría veritista.

Comencemos con la cláusula (3), que afirma que al incorporar un fenómeno desconcertante a un corpus cognitivo no se debe apelar a otro fenómeno igualmente desconcertante. Recordemos que la definición de *fenómeno desconcertante* se da en términos de probabilidad. En particular un fenómeno desconcertante F es uno que es altamente improbable, dado K , y para el que, además, se posee un argumento inductivo que tiene como conclusión que $\neg F$. Es decir, el requisito dice que *no* se puede incorporar un fenómeno desconcertante apelando a proposiciones que muy probablemente son falsas, es decir proposiciones que describen fenómenos que muy probablemente no son el caso. Al contrario, al incorporar fenómenos al corpus cognitivo, uno debe hacerlo mediante argumentos en los que la conclusión es altamente probable, dada la evidencia. En otras palabras, el requisito (3) nos dice que *no* debemos razonar de modo que las conclusiones de nuestros razonamientos *no*

ejemplifiquen la relación de ser altamente probables dada la evidencia. En otras palabras, el requisito nos dice que debemos razonar de modo tal que nuestras conclusiones ejemplifiquen esa relación, es decir que ejemplifiquen la relación de ser altamente probables, dada cierta evidencia. Así, también razonar de acuerdo a lo que dicta el requisito (3) ejemplifica esta relación adecuada con la creencia verdadera.

Para explicar en qué consiste el valor epistémico de las condiciones (1) y (2) que debe cumplir toda incorporación para arrojar entendimiento, hay que abordarlas juntas. Me parece que los autores las presentan como dos condiciones separadas sólo por cuestiones de análisis. Sin embargo, me gustaría argumentar que el valor de la no-circularidad (es decir, el valor de (1)) depende del valor de la unificación con los datos empíricos (es decir del valor de (2)). Quiero sostener que la no-circularidad de un corpus sólo es epistémicamente valiosa en la medida que conecte a los fenómenos de éste, en última instancia, con los datos. Argumento, por lo tanto, que el valor epistémico de la no circularidad depende del valor epistémico de aquéllos elementos básicos a los que toda conexión debe dirigirse en última instancia. Esto quiere decir que la no circularidad de un corpus cognitivo ejemplifica la relación de dependencia con la creencia verdadera, pues su valor epistémico depende del valor de los datos empíricos de un corpus cognitivo. Así, lo que se debe hacer es explicar el valor epistémico de tener creencias que describen datos empíricos en un corpus cognitivo dado. La explicación que quiero proponer es bastante sencilla: sostengo que las proposiciones que describen datos son valiosas epistémicamente porque están justificadas epistémicamente. Y, a su vez, el proceso de formar creencias que están epistémicamente justificadas ejemplifica la relación de ser conducente a la verdad.³⁶ La intuición de fondo que está detrás de esta propuesta es que la coherencia de un corpus cognitivo es valiosa sólo si los elementos que constituyen el corpus tienen, a su vez,

³⁶ Véase la sección (2.1.1). Al final de la sección (3.6) abordaré con más detalle la cuestión de si el valor de la justificación epistémica efectivamente puede explicarse en términos de una teoría veritista.

algún valor epistémico independiente de su coherencia con el corpus. Imaginemos que tenemos un sistema coherente que descansa, en última instancia, en un conjunto de creencias que sabemos que son falsas o bien para las que no tenemos ninguna justificación. Sin duda, la coherencia de un sistema como este no tendría ningún valor epistémico. Por otro lado, tener un corpus cognitivo cuyas creencias *no* están conectadas de un modo circular depende de que las creencias que están en su base estén epistémicamente justificadas. De modo que tener un corpus en el que no hay circularidad y en el que hay unificación con los datos empíricos ejemplifica la relación de dependencia con la creencia verdadera, pues según la definición de esta relación decimos que una unidad de evaluación ejemplifica la relación de *dependencia* con la creencia verdadera cuando su valor depende de que otras creencias (u otras actitudes dóxicas) sean verdaderas o bien tengan un valor epistémico distinto de la verdad que sea explicable en términos de la TVEE. En este caso, se ejemplifica la relación porque el valor del corpus cognitivo de este tipo *depende* de que otras creencias estén justificadas epistémicamente.

Esta explicación del valor epistémico de los requisitos (1) y (2) es consistente con la teoría de Schurz y Lambert. Ellos sostienen que la incorporación de cualquier dato empírico en un corpus cognitivo tiene una ganancia intrínseca. Es decir, la incorporación de un dato empírico en un corpus cognitivo incrementa, en principio, el grado de unificación del corpus cognitivo. Es posible explicar esta *ganancia intrínseca* de todo dato incorporado en términos de que todos los datos tienen un valor epistémico independiente: a saber, están justificados epistémicamente.

Así, todas las características necesarias para entender un fenómeno de acuerdo con esta teoría ejemplifican alguna de las relaciones con la creencia verdadera que presenté en la sección (1.2). Por lo tanto, el valor del entendimiento científico que surge de la incorporación de un fenómeno a un corpus cognitivo es también explicable en términos de una teoría veritista. A continuación analizaré si es posible explicar en términos del veritismo el valor epistémico del entendimiento científico que surge de modelos idealizados en la actividad científica.

3. ¿Es posible explicar en términos veritistas el logro epistémico asociado con uso de idealizaciones y de modelos en la ciencia?

Quizás el reto más grande que debe enfrentar el teórico veritista que desea explicar el entendimiento científico en términos de su teoría consiste en dar cuenta del valor de la forma de entendimiento científico asociada con el uso de idealizaciones y de modelos en la ciencia. Como veremos más adelante, algunos filósofos, y entre ellos Catherine Elgin de manera prominente, han insistido en que el logro epistémico que surge del uso de idealizaciones y modelos puede identificarse con una forma de entendimiento científico. Por ejemplo, Peter Lipton argumenta enérgicamente a favor de la idea de que el entendimiento puede alcanzarse mediante una multitud de herramientas, además de la explicación científica; por ejemplo,

mediante el uso de modelos científicos. Lipton concede que es tentador identificar al entendimiento con la explicación, pues explicar y entender están relacionados cercanamente. Según él, explicar es responder a preguntas-por-qué y entender es simplemente tener esas respuestas. Sin embargo, argumenta Lipton, podemos alcanzar entendimiento científico sin hacer uso de explicaciones. Por ejemplo, uno puede alcanzar entendimiento usando imágenes o modelos físicos (Cfr. Lipton, 2009, p. 45). Un ejemplo ilustra de un modo simple esta idea: uno puede entender en qué consiste el movimiento retrógrado de algunos planetas atendiendo a una demostración en un planetario con un modelo a escala. El modelo a escala no sólo proporciona entendimiento del fenómeno, sino que lo logra de un modo más directo y sencillo que una elaborada explicación. De acuerdo con los filósofos que defienden esta postura, no sólo los modelos a escala proporcionan entendimiento, sino que también lo hacen los modelos matemáticos.

Ahora bien, la tarea de este capítulo es doble: en primer lugar hay que identificar de una manera clara en qué consiste el logro epistémico que surge del uso de las idealizaciones y los modelos científicos. Como veremos a continuación, esta cuestión es bastante controvertida y no hay un consenso real sobre cuál es este logro. La respuesta que consiste simplemente en decir que el uso de modelos e idealizaciones nos proporcionan entendimiento científico no es realmente informativa, pues ya vimos que el entendimiento científico puede surgir de las explicaciones científicas y de la incorporación de un fenómeno a un corpus cognitivo. Es necesario que se explique con mayor detalle cuál es la naturaleza del logro epistémico que surge del uso de modelos, tal que sea posible identificarlo con una forma *distintiva* de entendimiento científico. Haciendo una revisión de la bibliografía al respecto buscaré dar una respuesta satisfactoria a esta pregunta. En segundo lugar, hay que investigar si es posible explicar el valor de esta forma de entendimiento científico en términos de la teoría veritista que he adoptado en esta investigación.

Dar cuenta de esta forma de entendimiento científico es quizás el reto más grande para el teórico veritista. Esto se debe a que una buena parte de los filósofos involucrados en el debate sobre la naturaleza de los modelos y las idealizaciones tienden a favorecer la tesis de que la veracidad no tiene un papel relevante en la explicación del valor de estas herramientas teóricas de representación. Pues, según ellos, los modelos y las idealizaciones científicas suelen ser representaciones estrictamente falsas. Pero, de acuerdo con estos filósofos, a pesar de su falsedad, estas herramientas tienen un gran valor en el desarrollo de la actividad científica: es un hecho evidente que el uso de los modelos y las idealizaciones permea una gran parte de la actividad científica. Si concedemos estos puntos, entonces el veritismo enfrenta un reto al explicar epistémicamente el valor de estas herramientas en sus propios términos, pues dada la supuesta falsedad que permea a este tipo de representaciones, al menos en primera instancia, no es claro cómo podría lograrlo. Eric Winsberg presenta el problema de la siguiente manera:

En una primera aproximación, las ficciones son representaciones que no se ocupan de la verdad. La ciencia está repleta de representaciones. Pero las representaciones que nos ofrece la ciencia [...] supuestamente buscan la verdad. [...] Si el objeto inmediato y apropiado de las ficciones es contrario a las metas de la ciencia, ¿qué papel podría haber para las ficciones en la ciencia? (Winsberg, 2009, p. 179).

Winsberg utiliza el término genérico “ficciones” para designar a las representaciones supuestamente no verídicas, entre ellas los modelos científicos.³⁷ Este problema general, que tiene que ver estrictamente con cuál es el logro epistémico asociado con el uso de representaciones falsas en la ciencia, tiene una ramificación que es de interés central para esta

³⁷ La terminología de “ficciones” tiene su antecedente en la filosofía de Hans Vaihinger, quien especialmente en su libro *The Philosophy of “As if”, A system of the Theoretical, Practical and Religious Fictions of Mankind* presenta la idea de que los hombres a menudo aceptan de manera consciente ficciones, ya sea para entender mejor el mundo mismo o para actuar de manera más eficiente en él.

investigación: ¿cómo explicar de un modo veritista el tipo de entendimiento científico que es producto de modelos e idealizaciones, i.e. ficciones científicas?

Pero antes de continuar con una discusión más profunda y detallada de esta cuestión es necesario hacer una serie de precisiones sobre la naturaleza de los modelos y las idealizaciones.

3.1. Consideraciones preliminares sobre los modelos y las idealizaciones

En primer lugar es de vital importancia trazar claramente una distinción entre las idealizaciones y los modelos científicos que a menudo se pasa por alto. En ocasiones se suele hablar de los modelos y las idealizaciones como si ambos pertenecieran al género de las ficciones científicas.³⁸ A pesar de que es posible encontrar algunas características que bien pueden justificar la inclusión de ambas dentro del mismo tipo de cosas (esto es, dentro del género de las ficciones), presentar así las cosas puede resultar confuso. La razón por la que me parece que se deben de mantener separadas la idealización y los modelos científicos consiste en que son cosas bien diferentes, en el sentido de que pertenecen a categorías ontológicas completamente distintas: los modelos son un tipo de *representaciones*, mientras que las idealizaciones son un tipo de *actividad*. Es cierto que también puede hablarse de la práctica de construcción de modelos, que es también una actividad y estaría, por lo tanto, en la misma categoría ontológica que la idealización. El error consiste en que a menudo no se hace explícita esta diferencia y se suele agrupar dentro de la misma categoría (la de las ficciones) a dos elementos que no pertenecen al mismo tipo ontológico. Quizás la razón por la que se suele caer en esta confusión consiste en que la construcción de modelos científicos regularmente implica algún tipo de idealización. Es importante, sin embargo, trazar claramente la distinción, no sólo para evitar cometer este error *categorial*, sino porque más adelante pondré en duda la idea de que las

³⁸ Esto es particularmente evidente en los ensayos que aparecen en *Fictions in Science. Philosophical Essays on Modeling and Idealizations* (2009), editado por Mauricio Suárez.

ficciones científicas sean falsas. De acuerdo con algunos filósofos que no hacen esta distinción, la clase de las *ficciones científicas* incluye tanto a los modelos como a la idealización. Ahora bien, es importante notar que las representaciones, pero no las actividades, son el tipo de cosas que son susceptibles de ser verdaderas o falsas. Por lo tanto, la tesis sobre la veracidad de las “ficciones” se verá reducida a la tesis sobre la veracidad de los modelos científicos. De acuerdo con la visión de las cosas que presento aquí, el papel de las idealizaciones (que son un tipo de actividad) consiste en generar, entre otras cosas, modelos científicos idealizados.

Una vez que ha quedado clara esta distinción es necesario decir más sobre qué son los modelos y las idealizaciones y es también necesario dar una clasificación general de ambos. De hecho, hay un gran consenso sobre qué son los modelos y las idealizaciones, pero hay muy poco consenso cuando se trata de dar una clasificación exhaustiva de ambos. Para los propósitos de esta investigación no es necesario entrar en muchos detalles sobre cuál es la manera óptima de clasificar a los tipos de modelos y de idealizaciones científicas. Asumiré aquí una clasificación sólo para presentar un panorama general de los modelos y las idealizaciones que se usan en la actividad científica; sin embargo asumir tal clasificación no es esencial para la discusión posterior, pues la postura que defenderé se sostiene incluso si se adoptara una clasificación alternativa. Sí es necesario, sin embargo, decir algo más sobre qué son los modelos y las idealizaciones.

Comencemos con los modelos. De acuerdo con la distinción que tracé al inicio de esta sección, todos los modelos científicos son representaciones. Sin embargo, de acuerdo con la tradición, hay al menos dos funciones representacionales fundamentales que puede cumplir un modelo, aunque aquí sólo me interesa *una* de éstas. La primera función (y que no me interesa) es la siguiente: un modelo representa a una teoría cuando da una interpretación de sus axiomas y teoremas. Este uso del término “modelo” tiene su origen en la lógica moderna: “un modelo es una estructura que hace a todas las oraciones de una teoría verdadera, donde una teoría se

interpreta como un conjunto (usualmente cerrado deductivamente) de oraciones en un lenguaje formal”³⁹ (Frigg y Hartmann, 2012, p. 12). Quizás un ejemplo puede aclarar esta noción:⁴⁰ decimos que cualquier estructura de la que sean verdaderos los axiomas y los teoremas de la geometría euclidiana es un modelo de ésta. Podemos elaborar, por ejemplo, un programa computacional en el que todos los axiomas y teoremas de la geometría euclidiana se cumplen; decimos, entonces, que el programa computacional es un modelo que representa adecuadamente a la teoría. Sin embargo, este sentido de “modelo” y este sentido de “representar” no son los más importantes para los intereses de esta investigación.

La segunda función representacional de los modelos tiene que ver mucho más con los propósitos de esta investigación. Esta función consiste en representar el mundo y hay al menos dos tipos de cosas que los modelos pueden representar: fenómenos o datos. Aquí sólo me enfocaré en los modelos que representan fenómenos. De cualquier manera explicaré en qué consiste cada uno.

Un modelo de datos es una versión idealizada de los datos que se obtienen por experimentación y observación. La utilidad de estos modelos surge a la hora de confirmar alguna teoría: regularmente se evalúa el poder predictivo de una teoría particular contrastándolo con modelos de datos, no con los así llamados “datos crudos” (*Cfr.* Frigg y Hartmann, 2012, p. 9). A pesar de que hay varias cuestiones metodológicas y filosóficas interesantes relacionadas con los modelos de datos y con su uso en la actividad científica, este tipo de modelos no resultan de interés central en esta investigación. Aun cuando estos modelos tienen una importancia sobresaliente en la actividad científica, como ya señalé antes, su valor está principalmente vinculado con la práctica de la confirmación de teorías. Aquí estoy interesado

³⁹ Las teorías semánticas sobre la naturaleza de las teorías científicas hacen un uso prominente de esta noción de modelo cuando ofrecen su caracterización de una teoría científica, de acuerdo con estas concepciones, las teorías científicas son conjuntos de modelos, presentar una teoría, nos dicen, es presentar una clase de modelos. Véase el capítulo 10 de *Fundamentos de teoría de la ciencia* (1997) de Díez y Moulines

⁴⁰ También el ejemplo lo tomo de Frigg y Hartmann (2012).

en los logros epistémicos asociados con el uso de modelos que pueden ser identificados con el entendimiento. A diferencia de los modelos de datos, los modelos que representan fenómenos sí han sido relacionados con el logro cognitivo del entendimiento científico. Por esta razón mi atención ahora se dirige a ese tipo de modelos.

Los tipos de modelos que abordaré ahora son aquellos que representan fenómenos del mundo real. Mi concepción de qué es un modelo de este tipo seguirá muy de cerca la caracterización de Michael Weisberg, quien sostiene que “los modelos son estructuras abstractas o físicas que pueden potencialmente representar fenómenos del mundo real” (Weisberg, 2007a, p. 216). Hay que especificar qué es un fenómeno para que esta definición sea más clara. Adoptaré la caracterización amplia de fenómeno que ofrecen Frigg y Hartmann; de acuerdo con ellos, “‘fenómeno’ se usa como un término sombrilla que cubre todas las características relativamente estables y generales del mundo que son interesantes desde un punto de vista científico” (Frigg y Hartmann, 2012, p. 3). Por ejemplo, la atracción gravitacional entre los cuerpos, la evolución por selección natural, la naturaleza de los enlaces químicos o la erupción de un volcán constituyen todos ellos fenómenos de interés científico que pueden, en principio, ser representados mediante un modelo científico.

Ahora bien, en lo que respecta a la clasificación de este tipo de modelos, comenzaré adoptando una clasificación muy general que también presenta Weisberg: los modelos pueden ser físicos o abstractos. Típicamente se asume que los modelos más relevantes para la actividad científica son los abstractos y, entre ellos, los modelos matemáticos son los más relevantes. Parte del problema de dar una clasificación más informativa de los modelos que representan fenómenos surge del hecho de que no hay siquiera una terminología común, de modo que cada autor parece dar una clasificación propia, que en ocasiones no tiene nada en común con la de

otros autores, aunque en ocasiones estén hablando de lo mismo.⁴¹ Aquí, por lo tanto, adoptaré sólo una clasificación mínima que parece ser la clasificación más general y que goza de mayor consenso.

La tendencia a la *representación idealizada* de los fenómenos es la característica más relevante de estos modelos, por lo tanto hay que explicar con mayor detalle en qué consiste la actividad de la idealización científica. Aquí, al igual que en el caso de los modelos, hay consenso en lo que respecta a una caracterización general, pero encontramos mucha disparidad cuando buscamos una clasificación de los tipos de idealizaciones que existen. En términos generales, la mayoría parece estar de acuerdo en que la idealización es una actividad que consiste en la distorsión deliberada de un objeto.⁴² Esta caracterización sugiere, evidentemente, que la idealización no es una actividad exclusiva de la actividad científica. Es importante, sin embargo, señalar que aquí sólo me enfocaré en la idealización propiamente científica. A pesar de que en ocasiones se habla de la idealización como si ésta fuera una entidad o un enunciado (por ejemplo, se suele decir que la ley de los gases ideales es una idealización), lo más correcto es hablar de ésta en términos de actividad. Esta actividad puede aplicarse, en principio, a una variedad de entidades científicas: modelos, teorías, leyes, diagramas, etc. De este modo, de

⁴¹ Podemos darnos cuenta de la magnitud del desacuerdo sobre una clasificación más fina de los modelos si echamos un vistazo al sumario que nos presentan Frigg y Hartmann de los tipos de modelos que aparecen en la literatura: modelos de sondeo, modelos fenomenológicos, modelos computacionales, modelos de desarrollo, modelos explicativos, modelos empobrecidos, modelos de prueba, modelos idealizados, modelos teóricos, modelos a escala, modelos heurísticos, modelos de caricatura, modelos didácticos, modelos de fantasía, modelos de juguete, modelos imaginarios, modelos matemáticos, modelos sustitutos, modelos icónicos, modelos formales, modelos analógicos y modelos instrumentales (Frigg y Hartmann, 2012, p. 1).

⁴² Por ejemplo, Roman Frigg sostiene que “una idealización es una simplificación deliberada de algo complicado con el objetivo de volverlo más tratable”. (Frigg y Hartmann, 2012). McMullin sostiene que “el término 'idealización' es bastante vago. Yo asumiré que significa una simplificación deliberada de algo complicado (una situación, un concepto, etc.) con la finalidad de alcanzar al menos un entendimiento parcial de esa cosa” (McMullin, 1985, p. 248). Michael Weisberg dice: “los filósofos de la ciencia reconocen cada vez más la importancia de la idealización: la introducción intencional de distorsión en las teorías científicas” (Weisberg, 2007b, p. 639). Por su parte, Moti Mizrahi sostiene que “en la literatura sobre modelos científicos, una idealización se toma como una simplificación deliberada de algo complicado con la finalidad de hacerlo más manejable como objeto de estudio” (Mizrahi, 2011, p. 5). Resulta impresionante que todos estos autores utilizan casi la misma frase para describir la idealización científica y, de hecho, varios de ellos inmediatamente nos presentan los mismos ejemplos para ilustrar su caracterización: los planos sin fricción, el modelo de bola de billar de los gases, entre otros. Esto parece sugerir efectivamente que hay un alto grado de consenso sobre el modo general de caracterizar a las idealizaciones.

acuerdo con esta terminología, sí tiene sentido hablar de modelos idealizados, leyes idealizadas, etc., pero no de idealizaciones sin más.

Ahora bien, a pesar del consenso sobre qué es una idealización, resulta más complicado dar una clasificación de las idealizaciones científicas, pues hay una amplia variedad de clasificaciones alternativas. Aquí adoptaré la clasificación que presentan Frigg y Hartmann, pues es bastante simple y precisamente por esta razón puede dar cabida a otras clasificaciones más finas. De acuerdo con ellos, podemos clasificar a las idealizaciones en de dos tipos: la aristotélica y la galileana:

Idealización aristotélica: consiste en remover todas las propiedades de un objeto o fenómeno concreto que creemos que no son relevantes para el problema en cuestión. Esto nos permite enfocarnos en un conjunto limitado de propiedades.

Idealización galileana: involucra la distorsión deliberada de las características de un objeto o fenómeno concreto. Esto permite volver a la representación del fenómeno que se estudia mucho más tratable computacionalmente (Cfr. Frigg y Hartmann, 2012).⁴³

En otras palabras, la idealización aristotélica consiste en ignorar algunas características del fenómeno, mientras que la idealización galileana consiste en distorsionar algunas características que el objeto sí tiene.

Centraré mi atención en los modelos idealizados (así sea en el sentido aristotélico o galileano), pues el papel que tienen en la actividad científica es bastante relevante y además representan un reto para el teórico veritista. Con esto no quiero decir que sean el único tipo de

⁴³ Como ya mencioné, la simplicidad de esta caracterización nos permite dar cabida o explicar clasificaciones alternativas. Por ejemplo, Michael Weisberg (2007b) propone la categoría de “idealización mínima”, la cual consiste en elaborar una representación que se limita a presentarnos los factores causales relevantes de un fenómeno. Esta idealización bien puede ser considerada un tipo de idealización aristotélica que tiene la particularidad de representar correctamente sólo los factores causalmente relevantes, mientras deja fuera todos los demás.

modelos ni el único tipo de representaciones importantes para la actividad científica, simplemente sostengo que son los más relevantes dados los fines de esta investigación.

A continuación presentaré un par de ejemplos para ilustrar la noción de modelo idealizado. Estos ejemplos serán útiles más adelante como casos de discusión. El primer modelo que quiero presentar es un modelo matemático que busca representar un sistema biológico en el que interactúan dos especies: una presa y un predador, conocido como las ecuaciones Lotka-Volterra.⁴⁴ De acuerdo con Michael Weisberg, después de la primera guerra mundial la población de ciertos tipos de peces (como los calamares, algunos tipos de bacalao y langosta) en el mar Adriático había disminuido, mientras que la población de tiburones, rayas y otros predadores había aumentado. Este hecho resultaba sorprendente, pues durante la guerra la pesca había disminuido considerablemente y se creía que esto daría una oportunidad para que la población de peces aumentara. Para responder a la pregunta de por qué la disminución de la pesca asociada con la guerra tuvo como resultado un incremento en la población de predadores y un decremento en la población de presas, Vito Volterra desarrolló un modelo matemático simple que representaba el sistema biológico del mar Adriático. Ahora bien, el modelo que presentó Volterra era un modelo idealizado. Para empezar, su modelo estaba compuesto sólo por una población de predadores y una de presas (ignorando que cada una de esas poblaciones estaba a su vez conformada por una gran variedad de especies) y les atribuyó pocas propiedades. Entre las propiedades que le atribuyó al sistema se encontraban un ritmo de crecimiento exponencial intrínseco para las presas en ausencia de predadores y un ritmo de muerte constante para los predadores. Sin embargo hay una gran cantidad de propiedades que el modelo no toma en cuenta y que sí se encontraban presentes en el ecosistema del mar Adriático: las propiedades espaciales del ecosistema, la densidad de las poblaciones, el clima y

⁴⁴ Tomo este ejemplo de “Who is a Modeler?” (2007a) de Weisberg y gran parte de lo que digo aquí sobre el modelo es una glosa de lo que presenta él en ese artículo.

el microclima o la interacción entre otras especies no incluidas dentro del grupo de presas o predadores. Así, el modelo de Volterra es un modelo idealizado del tipo aristotélico, pues prescinde de los elementos que no se consideran relevantes. A pesar de esto, el modelo efectivamente arrojó luz sobre el fenómeno, pues no sólo describió correctamente el comportamiento de las poblaciones de presas y predadores en el mar Adriático, sino que el modelo es capaz de predecir qué sucederá en sistemas biológicos que satisfacen situaciones similares a las del mar Adriático. Ahora bien, la primera meta de este capítulo consiste en identificar cuál es el logro cognitivo asociado con el uso de modelos idealizados. Es cierto que aquí ya se identificó un par de logros epistémicos, a saber la *descripción correcta* del comportamiento de las poblaciones de presas y predadores en ecosistemas similares al mar Adriático, y la *predicción* del comportamiento de estas poblaciones en ecosistemas similares al mar Adriático. Sin embargo, la cuestión fundamental consiste en explicar cómo se alcanzan estos logros epistémicos haciendo uso de modelos idealizados, esto es haciendo uso de representaciones supuestamente falsas. En otras palabras, la pregunta relevante es: ¿cómo una representación distorsionada del mundo nos permite describir correctamente y predecir el comportamiento del fenómeno que representa? Es importante señalar que cualquier postura sobre cuál es el logro epistémico asociado al uso de modelos idealizados debe dar una respuesta satisfactoria a esta pregunta, entre otras cosas. En secciones posteriores propondré respuestas particulares a estas cuestiones. Por ahora continuaré con la presentación de ejemplos de modelos idealizados.

Pasemos a otro ejemplo: la ley de los gases ideales,⁴⁵ una ley que pretende predecir el comportamiento de los gases bajo ciertas condiciones. De acuerdo con esta ley, el estado

⁴⁵ Gran parte de la presentación de la ley de los gases ideales que hago aquí sigue de cerca la presentación que hace Moti Mizrahi en “Idealizations and Scientific Understanding”, de próxima aparición.

termodinámico de un gas está determinado por su presión, su volumen y su temperatura y se formula del siguiente modo:

$$pV = nRT$$

Donde n es el número de moles, R es la constante universal de los gases ideales, p es la presión, V es el volumen y T es la temperatura. Es importante señalar que esta ley no toma en cuenta el volumen que ocupan las moléculas del gas ni las fuerzas de atracción y repulsión entre ellas y, además, se asume que las colisiones entre las moléculas son perfectamente elásticas, pues toda la energía se expresa como energía cinética translacional. De acuerdo con Moti Mizrahi, para derivar la ley de los gases ideales hace falta aceptar los siguientes supuestos:

(GL1) Que las moléculas del gas se mueven en direcciones aleatorias. Si, por alguna razón, las moléculas del gas se movieran todas estrictamente hacia arriba o hacia abajo [...] en el contenedor, entonces la ley no se sostendría.

(GL2) Las interacciones entre las moléculas del gas pueden ser ignoradas. [...] En la mayoría de los casos, la distancia entre las moléculas de una muestra de gas es lo suficientemente grande como para ser ignorada en los cálculos.

(GL3) La energía promedio de las moléculas del gas es proporcional a la temperatura del gas (Mizrahi, 2011, p. 6).

Mizrahi llama a (GL1-GL3) *suposiciones idealizantes* precisamente porque describen de manera idealizada a los gases y su comportamiento. De hecho, estas suposiciones son estrictamente falsas: describen de manera incorrecta la realidad. Sin embargo, a pesar de que la ley de los gases ideales surja de la actividad de la idealización resulta que tiene un gran éxito predictivo en una gran variedad de circunstancias, de manera que esta ley también parece describir correctamente el comportamiento de los gases.

Parece claro que en los dos casos hay algo bueno desde el punto de vista epistémico asociado con el uso de modelos idealizados, en particular ya se han identificado dos logros: su *éxito predictivo* y su capacidad de *describir correctamente* el comportamiento de los fenómenos sobre los que tratan. En ambos casos los modelos nos ponen en una situación epistémica buena con respecto a los fenómenos que se estudian: nos permiten predecir y describir correctamente. La pregunta relevante que hay que intentar responder ahora es cómo se alcanzan estos logros epistémicos utilizando herramientas que supuestamente son representaciones no verídicas de los fenómenos estudiados.

3.2. Una propuesta insatisfactoria. La importancia del principio de que no haya premisas falsas.

Algunos filósofos de la ciencia han sugerido que el logro epistémico asociado con el uso de modelos idealizados es el conocimiento, al mismo tiempo que sostienen que los modelos idealizados son representaciones no verídicas del mundo. Me gustaría señalar que adoptar esta respuesta equivale a ignorar la discusión relevante y ampliamente aceptada en la epistemología contemporánea sobre el principio de que no haya premisas falsas [*no false lemmas principle*]. Llevar a cabo esta discusión tiene una importancia adicional, pues nos servirá para mostrar qué alternativas quedan disponibles para explicar cuál es el logro epistémico asociado con el uso de modelos idealizados.

Anouk Barberousse y Pascal Ludwig son unos de los pocos que defienden de manera explícita la postura que consiste en sostener tanto que los modelos idealizados proporcionan conocimiento, como que éstos son representaciones no verídicas:

Nuestra tesis principal es que los modelos representan situaciones ficticias, a saber, situaciones que no se dan y que no se pueden dar (dado lo que

sabemos) en el mundo. En otras palabras, los modelos son *ficciones*, esto es, precisamente, representaciones de situaciones ficticias. Más aún, representar situaciones ficticias les da el poder de transmitir nuevo conocimiento científico útil (Barberousse y Ludwig, 2009, p. 57).

En esta formulación de su postura, Barberousse y Ludwig dejan claro que se comprometen con los dos elementos que he identificado: 1) los modelos idealizados son representaciones no verídicas y 2) los modelos idealizados proporcionan conocimiento del mundo. Es importante resaltar que su postura con respecto a la no-veracidad de los modelos idealizados es bastante fuerte: los modelos idealizados no sólo representan hechos que no se dan, sino que representan hechos que *no se pueden* dar. Esto significa que, desde su punto de vista, los modelos idealizados son *necesariamente* representaciones no verídicas. Sin embargo, la crítica que elaboraré a continuación hacia las posturas de este tipo no depende de que se sostenga esta tesis más fuerte. Es decir, mi crítica aplica también a las posturas que sostengan una versión más débil de (1), en particular una que no se comprometa con que la no-veracidad de los modelos científicos sea necesaria. Para entender la crítica que dirigiré a estas posturas es necesario explicar unas cuestiones relacionadas con la justificación epistémica y el análisis del conocimiento proposicional.

El famoso artículo de Edmund Gettier “Is Justified true Belief Knowledge”, publicado en 1963, mostró, según consenso general, que el análisis tradicional del conocimiento proposicional que identifica a éste con la creencia verdadera y justificada, es insuficiente. Muchas teorías epistemológicas han surgido como intentos por resolver el problema que Gettier presenta en ese artículo. Por ejemplo, las teorías externistas del conocimiento (cuya importancia en la actualidad es enorme) tienen su antecedente en la teoría causal del conocimiento de Goldman, que constituye una respuesta temprana al desafío que representa el problema de Gettier.

Ahora bien, el principio de que no haya premisas falsas se propuso precisamente en el contexto de la discusión de este artículo. Y a pesar de que el consenso en la epistemología contemporánea nos dice que apelar a alguna versión de este principio no es suficiente para resolver el problema de Gettier, también es cierto que la corrección del principio como una condición necesaria para la justificación epistémica también es ampliamente aceptada en la discusión contemporánea. Debemos recordar que el problema que presenta Gettier en su artículo consiste en desafiar el análisis tradicional del conocimiento (que identifica a éste con la creencia verdadera y justificada) presentándonos un par de contraejemplos que muestran que las condiciones del análisis tradicional no son suficientes para saber que p . Lo que muestran los ejemplos tipo Gettier es que el conocimiento es incompatible con un tipo de suerte epistémica. Uno no puede adquirir conocimiento, aunque su creencia sea verdadera y esté justificada si llega a esa creencia verdadera de un modo meramente accidental. El sujeto no se da cuenta de que su justificación es defectuosa y que su creencia es verdadera por una razón que es completamente independiente de su justificación. Ahora bien, el par de ejemplos que presenta Gettier en su artículo presentan casos en los que los sujetos utilizan en su proceso de razonamiento una creencia falsa, pero justificada. Una de las estrategias iniciales para superar el problema de Gettier consistía en señalar que esos casos particulares no constituyen un contraejemplo para el análisis tradicional, basándose en la idea de que violan un principio que debe cumplirse siempre que alguien tiene conocimiento proposicional: el principio de que no haya premisas falsas. Este principio ha sido presentado de distintas maneras, pues ha sido reformulado en varias ocasiones para fortalecerlo, dependiendo de las objeciones que ha enfrentado. Aquí no hace falta explorar detalladamente las diversas formulaciones del principio que se han propuesto, simplemente adoptaré la siguiente formulación que goza de un amplio consenso y resulta bastante inteligible:

Si h está justificado para S con base en la evidencia e , entonces ningún elemento de e que sea *esencial* para esa justificación es falso. Una proposición p , es esencial en este contexto, digamos, sólo si S no estaría justificado en aceptar h si se eliminara p de su evidencia, y no ocurrieran otros cambios (Pappas y Swain, 1978, p. 14).

Ahora bien, la razón por la que se considera que añadir este principio al análisis tradicional del conocimiento no basta para superar el problema de Gettier consiste en que se pueden elaborar nuevos contraejemplos que sí satisfacen el principio de que no haya premisas falsas y que muestran que el análisis enmendado es insuficiente.⁴⁶ Así, el hecho de que se puedan elaborar estos contraejemplos ha llevado a la conclusión de que este principio no es la clave para resolver el problema en el análisis tradicional del conocimiento que presenta Gettier. Sin embargo, también es cierto que se asume que este principio es correcto y que uno no puede adquirir conocimiento si éste se justifica apelando de un modo esencial a alguna premisa falsa.

Ahora bien, la postura ejemplificada por Barberousse y Ludwig que sostiene que el logro epistémico de los modelos idealizados consiste en proporcionar conocimiento del mundo es inaceptable, pues representa una clara violación del principio ampliamente aceptado de que no haya premisas falsas. Hay que recordar que esta postura se compromete con la tesis de que los modelos idealizados son representaciones no verídicas del mundo. Esto significa que, de acuerdo con ellos, uno puede utilizar representaciones no verídicas en un proceso de razonamiento, para alcanzar conocimiento. Pero el principio de que no haya premisas falsas sostiene que la justificación que es parte del conocimiento no puede descansar en premisas falsas. Es cierto que el uso de modelos idealizados nos pone en una situación epistémicamente

⁴⁶ El capítulo 3 de Shope (1983) y la introducción de Pappas y Swain (1978) presentan de manera esquemática la dialéctica entre quienes defienden que la inclusión del principio de que no haya premisas falsas resuelve el problema y los defensores de la idea de que el problema de Gettier se mantiene a pesar de la inclusión de este requisito. Ahí se pueden encontrar las distintas formulaciones del principio, así como las razones por las que se considera que cada una de esas versiones no es útil para solucionar el problema de Gettier.

favorable con respecto a los fenómenos que éstos representan. Sin embargo, Barberousse y Ludwig cometen el error de identificar apresuradamente a este logro epistémico con el conocimiento.

En este punto es de vital importancia señalar que el problema con este tipo de postura no consiste meramente en que identifique al logro epistémico asociado con el uso de modelos idealizados con el conocimiento, sino que el problema consiste en que hay una inconsistencia dentro de su misma postura: no se puede asumir que los modelos idealizados son representaciones no verídicas al mismo tiempo que se asume que es posible utilizarlas para alcanzar conocimiento. La inconsistencia entre estas dos tesis se vuelve evidente cuando se muestra que su combinación lleva a la violación del principio de que no haya premisas falsas.

En otras palabras, es inconsistente la triada que forman las dos tesis presentes en la postura de Barberousse y Ludwig y el principio de que no haya premisas falsas. Estas tesis son 1) que los modelos idealizados son representaciones no verídicas del mundo y 2) que los modelos idealizados proporcionan conocimiento del mundo. En esta investigación asumiré que el principio de que no haya premisas falsas es correcto, así que buscaré eliminar la inconsistencia en esta triada haciendo algún cambio a una de estas tesis. Estas consideraciones sugieren, por lo tanto, que si queremos respetar el principio de que no haya premisas falsas, quedan disponibles al menos dos vías alternativas para explicar en qué consiste el logro epistémico asociado con los modelos idealizados:

- a) una estrategia que sostenga que los modelos idealizados son efectivamente representaciones no verídicas del mundo, pero que identifique al logro epistémico asociado con ellos con uno diferente del conocimiento, en particular con un logro no-fáctico, es decir un logro que no implica la veracidad de la representación anidada en el logro epistémico y

b) una estrategia que rechaza que los modelos idealizados efectivamente sean necesariamente representaciones no verídicas y que identifique al logro cognitivo asociado con ellos con el conocimiento o con otro logro epistémico fáctico.

Cualquiera de estas dos posiciones es consistente con el principio de que no haya premisas falsas. La estrategia (a) acepta que los modelos idealizados son representaciones no verídicas pero insiste en que éstos *no* proporcionan conocimiento del mundo. Esto significa que no habrá casos de conocimiento en los que la justificación esté formada por representaciones no verídicas. Por lo tanto, la postura es consistente con el principio. Por otro lado, la estrategia (b) rechaza que los modelos idealizados son representaciones no verídicas y sostiene que uno puede alcanzar conocimiento utilizándolos. Como los modelos pueden ser representaciones verídicas, su uso para alcanzar conocimiento no implica una violación del principio de que no haya premisas falsas. Por lo tanto, la estrategia (b) también es compatible con este principio. Lo que no podemos hacer coherentemente es sostener el principio de que no haya premisas falsas y además que los modelos idealizados son representaciones no verídicas que podemos utilizar para alcanzar conocimiento del mundo.

En la siguiente sección presentaré la propuesta de Catherine Elgin, que es una postura tipo (a), y argumentaré que no es una propuesta satisfactoria. Posteriormente presentaré la solución que me parece más adecuada y que es una postura tipo (b). Pero antes me gustaría señalar algunas precisiones que deben hacerse a la argumentación que aquí he presentado en contra de la postura ejemplificada por Barberousse y Ludwig.

He argumentado que los supuestos de esta estrategia son inconsistentes con el principio de que no haya premisas falsas. Como ya vimos, el principio sostiene que una creencia verdadera no puede constituir conocimiento si la justificación de ésta descansa esencialmente en alguna premisa falsa. Sin embargo, es necesario hacer algunas precisiones para que esta línea

argumental efectivamente funcione. En primer lugar hay que aclarar cuál es la actitud dóxica involucrada cuando utilizamos modelos idealizados para alcanzar, supuestamente, conocimiento. Recordemos que Barberousse y Ludwig están comprometidos con la idea de que los modelos idealizados son representaciones no verídicas del mundo. Si esto es así, entonces resulta que ellos no pueden sostener que la actitud dóxica que tenemos con respecto a los modelos idealizados sea la *creencia*. Para entender la razón de esta imposibilidad hay que explicar brevemente algunas cuestiones sobre la naturaleza de la actitud dóxica de creer.

Al menos en la tradición filosófica analítica parece haber un acuerdo amplio en que, en términos generales, las creencias son actitudes disposicionales hacia una proposición o hacia el estado de cosas que ésta describe.⁴⁷ Por ejemplo, “creer que p es tomar como verdadero que p - tomar como verdadero que el estado de cosas descrito por “ p ” se da” (Chignell, 2010, p 15). Cohen define a la creencia de un modo similar: “la creencia de que p es una disposición a sentir normalmente que es verdad que p y falso que $no-p$, cuando uno pone atención a cuestiones que surgen con respecto a la proposición p , o a elementos que se refieren a ella, independientemente de si uno está dispuesto a actuar, hablar o razonar en consecuencia” (Cohen, 1992, p. 4). Es importante darse cuenta de la importancia que tiene la verdad en esta actitud dóxica. Si bien es cierto que hay mucho apoyo a favor de la tesis de que las creencias son disposiciones involuntarias,⁴⁸ parece imposible que uno pueda creer algo, de lo que esté plenamente consciente que no es verdadero. Tanto Velleman como Cohen sostienen que las creencias (a pesar de ser involuntarias) deben de estar abiertas a consideraciones evidenciales. Por ejemplo, Velleman sostiene que “creer involucra ver a una proposición como verdadera con la finalidad de verla así sólo si realmente es verdadera. De modo que creer una proposición es aceptarla con la finalidad de aceptarla como una verdad” (Velleman, 2010, p. 251). Es decir, la meta de

⁴⁷ Véase, por ejemplo, Andrew Chignell (2010), sección 3 (pp. 15-19); David Velleman (2000), y Jonathan Cohen (1992).

⁴⁸ Véase la sección I, §4, de Cohen (1992).

formar creencias consiste en formar creencias verdaderas. A manera de contraste, la suposición es una actitud dóxica en la que también se ve a p como verdadera, sin embargo la finalidad de esta actitud no tiene que ver con aceptar proposiciones verdaderas. La finalidad principal de la suposición consiste en saber que se sigue de razonar *como si p fuera verdadera*. De modo que la relación estrecha entre la creencia y la verdad da sustento a la idea de que la actitud dóxica involucrada en el uso de modelos idealizados no puede ser la de la creencia: uno no puede *creer* las proposiciones que articulan un modelo, pues uno *sabe* que tales proposiciones son *falsas*.

Es una característica común de los modelos idealizados que la idealización es completamente intencional, es decir, los científicos que los usan son completamente conscientes de la simplificación presente en los modelos que usan. Por esta razón decimos que la idealización involucrada en los modelos idealizados es una “distorsión intencional”. Esto significa, desde el punto de vista de Barberousse y Ludwig, que los científicos que hacen uso de modelos idealizados para alcanzar conocimiento deben ser plenamente conscientes de que utilizan premisas que incluyen representaciones no verídicas. Y, por lo tanto, dada la discusión anterior no podrían creerlas. Así, si uno no puede creer las proposiciones que articulan un modelo idealizado, entonces ¿cuál es la actitud dóxica apropiada involucrada en el uso de ellos?

Recordemos que mi argumentación en contra de la postura de Barberousse y Ludwig se basaba en señalar que, de acuerdo con sus supuestos, el conocimiento adquirido mediante el uso de modelos idealizados supone el uso de premisas falsas. El análisis que he hecho sobre este punto ha mostrado que las premisas falsas no pueden identificarse con creencias falsas. Si queremos que nuestra argumentación en contra de esta postura se sostenga debemos identificar cuál es la postura dóxica pertinente cuando se hace uso de modelos idealizados.

El candidato más fuerte que podría tener el papel que se busca aquí es la actitud de *aceptación*. Jonathan Cohen la define de la siguiente manera:

Aceptar la proposición o regla de inferencia p es tratarla como si fuera el caso que p . Más precisamente, aceptar que p es tener o adoptar una política que consiste en considerar, proponer o postular que p -i.e. Que consiste en incluir esa proposición o regla de inferencia entre las premisas que uno tiene para decidir qué hacer o pensar en un contexto particular, sea o no que uno sienta que es verdadero que p . (Cohen, 1992, p. 4).

Esta actitud dóxica efectivamente puede ser útil para caracterizar lo que sucede cuando usamos modelos idealizados y razonamos a partir de ellos para comprender mejor algún fenómeno natural. De acuerdo con Cohen, a pesar de que uno no crea que p , si uno decide aceptarla, puede utilizar a p como premisa en sus deliberaciones y razonamientos. Esto se debe, en parte, a que la aceptación, a diferencia de la creencia, sí es una actitud voluntaria. Además, el propio Cohen sostiene que las razones para aceptar p no tienen por qué ser evidenciales, sino que también pueden ser éticas, profesionales, prudenciales, religiosas, estéticas o de otro tipo (*Cfr.* Cohen, 1992, p. 20). Por ejemplo, a pesar de que uno no pueda creer que su ídolo literario ha cometido plagio, quizás la evidencia es tan apabullante que decide aceptarlo. Si esto es así, entonces es posible trazar el siguiente paralelo: a pesar de que los científicos sean conscientes de la idealización de las representaciones que utilizan, esto no les impide aceptarlas y utilizarlas como premisas en sus razonamientos.

Así, la aceptación parece ser el candidato más viable para caracterizar a la actitud dóxica que un sujeto tiene hacia los modelos idealizados. Si esta identificación es correcta, entonces la argumentación que presenté en un principio en contra de la postura de Barberousse y Ludwig se mantiene firme, simplemente hay que notar que la actitud dóxica que se tiene hacia las proposiciones que articulan al modelo idealizado es la de aceptación y no la de creencia. A pesar de esta precisión, su postura los lleva a suponer que las premisas que se utilizan cuando se hace uso de un modelo idealizado están constituidas por aceptaciones que son estrictamente no verídicas y, por lo tanto, están comprometidos a sostener que esas aceptaciones no verídicas

figuran en la justificación de la creencia verdadera y, así, violan el principio de que no haya premisas falsas.

Parece sensato, por lo tanto, explorar las alternativas que se presentaron anteriormente y que son consistentes con el principio de que no haya premisas falsas para investigar si alguna de ellas puede dar una respuesta satisfactoria a la pregunta de cuál es el logro epistémico asociado con el uso de modelos idealizados. Comenzaremos explorando la propuesta de Catherine Elgin que cae bajo el tipo de estrategia (a), que, recordemos, consiste en mantener el supuesto de que los modelos idealizados son representaciones no verídicas, pero identifica al logro cognitivo asociado con ellos con uno distinto del conocimiento.

3.3. La propuesta de Catherine Elgin. La interpretación no fáctica del entendimiento científico.

Elgin sostiene que el conocimiento es fáctico, esto quiere decir que uno sabe p sólo si p es verdadera. De aquí en adelante adoptaré su terminología y diré que un logro epistémico es fáctico siempre y cuando el logro epistémico en cuestión esté constituido por alguna representación verídica. Elgin también afirma (2007, p. 33) que, al igual que el conocimiento, el entendimiento científico es un tipo de logro epistémico. Sin embargo, sostiene que adoptar una concepción fáctica del entendimiento científico resultaría demasiado restrictivo. Antes de continuar con la exposición de la argumentación de Elgin hay que señalar rápidamente que ella sostiene que el objeto del entendimiento es un cuerpo de información comprensivo y no proposiciones individuales.⁴⁹ Es decir, según su postura, uno sólo puede tener entendimiento de un conjunto de información o de proposiciones. Uno no puede tener entendimiento sin más de un elemento informativo particular o de una proposición aislada. El entendimiento de una proposición simple deriva del entendimiento que el sujeto tiene de un conjunto de información.

⁴⁹ Esta concepción del entendimiento goza de un gran consenso entre los epistemólogos del entendimiento. Véase especialmente Kvanvig (2003), Zagzebski (2001), Grimm (2006) y Pritchard (2009).

Ahora bien, Elgin identifica dos cuestiones que impiden la adopción de una noción fáctica del entendimiento científico. El primero tiene que ver con la práctica cotidiana de atribución de entendimiento y el segundo tiene que ver con el uso de modelos idealizados en la ciencia. Podría parecer que, dados los fines de esta investigación, sólo el segundo problema tiene relevancia; sin embargo, como veremos a continuación, incluso el primer problema tiene una vertiente que tiene aplicación en contextos científicos. El primer problema consiste en que en contextos cotidianos somos perfectamente capaces de atribuirle a alguien entendimiento de algún cuerpo de información a pesar de que una parte de la información que entiende no sea verídica e incluso cuando algunos de los elementos centrales para su entendimiento no sean verídicos. El ejemplo paradigmático que presenta Elgin es el de un niño pequeño que tiene una visión bastante simplificada de, por ejemplo, la teoría de la evolución. Imaginemos que su entendimiento de la evolución humana incluye como una tesis central que los humanos descienden de los monos. Es cierto que este niño no tiene el entendimiento más completo que se puede tener del asunto, sin embargo también es cierto que su situación cognitiva es mejor que la del niño que piensa que los humanos no evolucionaron en absoluto o de alguno que piensa que los humanos descienden de las mariposas. Es necesario que la epistemología explique en qué consiste ese logro epistémico del niño. Elgin argumentará a favor de la idea de que este logro es efectivamente una clase de entendimiento. Pero, dada la presencia de falsedades en el cuerpo de información que se entiende, argumenta Elgin, debemos interpretar este logro epistémico como si fuera no fáctico.

Es importante señalar que este fenómeno no es exclusivo del contexto del aprendizaje, sino que también suele ejemplificarse en el desarrollo mismo de la ciencia, es decir “el patrón desplegado por el estudiante mientras se mueve de la postura ingenua de la evolución humana hasta alcanzar la postura que sostiene el profesor de la biología evolutiva es el mismo patrón que la ciencia despliega en la secuencia de teorías que desarrolla” (Elgin, 2007, p. 37). Es decir,

también en el desarrollo histórico de la ciencia nos encontramos con momentos en los que incluso el entendimiento de los científicos de vanguardia contiene elementos no verídicos y, sin embargo, parecería sensato decir que también ellos tuvieron entendimiento científico de los fenómenos que estudiaban.⁵⁰ Si esto es cierto, entonces la necesidad de una respuesta se vuelve más acuciante para el filósofo de la ciencia, pues debe explicar en qué consiste el logro epistémico alcanzado por los teóricos en etapas primitivas del desarrollo científico. A continuación tenemos un ejemplo que presenta la propia Elgin y que ilustra cómo el desarrollo científico de hecho presenta este patrón:

Un principio central de la teoría de Copérnico es que la Tierra viaja alrededor del sol en una órbita circular. Kepler mejoró a Copérnico afirmando que la órbita de la tierra no es circular, sino elíptica. Habiendo abandonado el compromiso con el espacio absoluto, los astrónomos actuales no pueden decir que la tierra viaja alrededor del sol *simpliciter*, sino que deben hablar de cómo la tierra y el sol se mueven cada uno en relación al otro. A pesar de que la afirmación central de Copérnico era estrictamente falsa, la teoría a la que pertenecía constituía un avance mayor en el entendimiento con respecto a la teoría Ptolemaica que reemplazó (Elgin, 2007, pp. 37-38).

Incluso en estos casos Elgin insiste en que el logro epistémico de estos científicos es el entendimiento. Así, tanto Ptolomeo, como Copérnico y Kepler tuvieron un grado de entendimiento científico sobre los movimientos planetarios, a pesar de que el cuerpo de información pertinente que cada uno tenía a su alcance incluía proposiciones centrales literalmente falsas. Ahora bien, es cierto que Kepler tuvo un mejor entendimiento que Copérnico y que éste, a su vez, tuvo un mejor entendimiento que Ptolomeo de los movimientos

⁵⁰ De hecho, dado el carácter refutable de la mayoría de las teorías científicas, sería sensato pensar que incluso los científicos contemporáneos tienen un entendimiento científico que depende de un modo esencial de representaciones no verídicas.

planetarios. Sin embargo, Elgin no ofrece una explicación de cuál es la diferencia entre el valor de los logros epistémicos de estos científicos. Más adelante, cuando presente mi propuesta sobre cómo caracterizar esta forma de entendimiento científico que se alcanza mediante el uso de modelos idealizados, explicaré cómo mi postura sí puede dar cuenta del fenómeno intuitivo de que el entendimiento se da en grados.

El segundo problema que identifica Elgin para una elucidación fáctica del entendimiento científico tiene que ver con una práctica muy extendida en la actividad científica: la idealización. Más arriba presenté el caso de la ley de los gases ideales: un modelo idealizado sobre el modo en que se comportan los gases. Elgin apoya gran parte de su argumentación en el análisis de este modelo. De acuerdo con ella, este modelo representa a los gases y su comportamiento de un modo literalmente falso: al intentar describir el comportamiento de los gases de un modo adecuado se asume que el gas está compuesto por moléculas perfectamente elásticas que no están sujetas a la fricción y que no tienen ningún tipo de atracción intermolecular. En este caso, el modelo idealizado se desvía de la realidad al menos en estos aspectos, lo cual da como resultado una representación estrictamente falsa de los gases. Elgin acertadamente señala que el uso de idealizaciones es muy extendido en la ciencia. Sin embargo sostiene algo más fuerte: que no hay ninguna expectativa de eliminar la idealización de las teorías científicas, que la “eliminación de idealizaciones no es un desiderátum” (Elgin, 2007, p. 38). Me parece que esta última afirmación es demasiado arriesgada y de hecho hay buenas razones para creer que es incorrecta, como argumentaré a continuación.

Es un hecho ampliamente aceptado que, en ocasiones, los científicos optan por elaborar una descripción idealizada de un fenómeno simplemente porque resulta computacionalmente imposible llevar a cabo cálculos que involucren alguna representación más completa del mismo. Michael Weisberg sostiene que la idealización que él identifica como “galileana” es la práctica de “introducir distorsión en las teorías con la meta de simplificarlas para hacerlas más

tratables computacionalmente” (Weisberg, 2007b, p. 640) y más adelante concede que la justificación de este tipo de idealización es pragmática y que está ligada con la tratabilidad computacional. Además, sostiene que “el avance en el poder computacional y técnica matemática debería llevar al idealizador galileano a des-idealizar, removiendo la distorsión y añadiendo de nuevo el detalle a sus teorías” (Weisberg, ídem, p. 641).⁵¹ Incluso podemos apelar al ejemplo paradigmático de Elgin, el de los gases ideales, para mostrar cómo efectivamente hay una tendencia a la des-idealización. Dado que esta ley ignora la atracción entre las moléculas y su tamaño, la ley es particularmente precisa cuando se aplica a gases monoatómicos a altas temperaturas y presiones bajas: en la medida que el gas ocupe un volumen mayor y la presión sea más baja, ignorar la atracción molecular y el tamaño de las moléculas resulta irrelevante: las moléculas están tan distantes entre sí que ignorar estas características no afecta en gran medida el resultado del cálculo. En la medida en que el gas se comprime y ocupa un volumen menor (esto es, en la medida que la densidad del gas aumenta) la ley de los gases ideales deja de dar resultados satisfactorios: “los gases reales no son tan compresibles a altas presiones como los gases ideales. El volumen de un gas real es mayor que lo que predice la ley de los gases ideales a altas temperaturas” (Mizrahi, 2011, p. 7). Por esta razón Johannes Diderik van der Waals propuso introducir una constante b en la ley de los gases ideales que representara el volumen que de hecho ocupan las partículas del gas. Así, a bajas densidades nb ⁵² sería demasiado pequeña para hacer una diferencia en el cálculo, pero en la medida que la densidad aumenta, nb da cuenta de la desviación y arroja resultados más precisos en este tipo de circunstancias. Parece que en este caso efectivamente encontramos una tendencia a la des-idealización, pues se añade un elemento que antes había sido ignorado: el volumen que ocupan las moléculas del gas. Esto muestra que la afirmación de Elgin según la

⁵¹ McMullin (1985, p. 261) está de acuerdo con esto; él sostiene que “los modelos pueden hacerse más específicos eliminando asunciones simplificadoras y 'des-idealizando”.

⁵² Donde n está por el número de moles del gas en cuestión.

cual la ciencia no debería buscar eliminar las idealizaciones científicas es al menos un poco arriesgada, pues es fácil encontrar casos particulares que sugieren lo contrario y apoyo teórico en contra de esa afirmación.

Sin embargo, esto no debe representar un gran problema para Elgin. Para que su argumento se sostenga basta con que las idealizaciones jueguen un papel relevante en la actividad científica, independientemente de si su eliminación es un desiderátum. Es cierto que el uso de modelos idealizados tiene un papel muy relevante en la actividad científica, y Elgin debería argumentar que eso basta para mostrarnos que la elucidación correcta del tipo de entendimiento que se alcanza mediante el uso de modelos idealizados es una no-fáctica.

Resumiendo, desde el punto de vista de Elgin, la interpretación no fáctica de este tipo de entendimiento supuestamente se ve favorecida por tres razones: es la única manera de explicar el papel que tienen los modelos idealizados en esta forma de entendimiento científico; es la única manera de dar cuenta del patrón que se exhibe en el desarrollo histórico de la ciencia, y por último, nuestro uso cotidiano del término es compatible con una interpretación no fáctica del entendimiento. Sin embargo hace falta exponer con detalle en qué consiste, para Elgin, el entendimiento científico y cuál es su valor epistémico.

Para Elgin, el entendimiento consiste en “una aprehensión de un cuerpo comprensivo de información que está basado en los hechos, que responde debidamente a la evidencia, que permite llevar a cabo inferencias y argumentos no triviales, y quizás permite llevar a cabo acciones que tienen que ver con el tema sobre el que trata la información” (Elgin, 2007, p. 39). Así, este logro epistémico consiste precisamente en aprehender el cuerpo de información apropiado y darse cuenta de las relaciones que hay dentro de él e incluso ser capaz de apreciar su posible utilidad teórica y práctica. Sin duda resulta sorprendente que Elgin adopte esta interpretación del entendimiento. Ella misma ha sostenido que una interpretación adecuada del término debe de servir para explicar el entendimiento rudimentario del niño que tiene una

visión altamente simplificada de la teoría de la evolución. Este ejemplo parece sugerir que el entendimiento debe ser bastante asequible, tan asequible que incluso un niño pequeño lo puede alcanzar. Sin embargo, Elgin define el entendimiento de un modo muy sofisticado, de un modo tal que termina dando una elucidación hiper-intelecualizada de qué es entender: parece muy cuestionable que un niño como el que presenta la propia Elgin efectivamente tenga una aprehensión de un cuerpo de información que satisfaga todos los requisitos que ella misma propone. Pero pasemos por alto este problema y analicemos con mayor detalle si esta elucidación del logro epistémico asociado con el uso de modelos idealizados resulta aceptable.

En este momento hace falta que Elgin dé una explicación de cómo los modelos idealizados proporcionan el tipo de entendimiento científico que ella misma ha identificado. Elgin da una respuesta explícita de cómo logran esto: “sugiero que las idealizaciones efectivas son falsedades felices. Nada en el mundo responde a ellas exactamente, así que como descripciones son falsas. Pero son felices en el sentido de que proveen acceso epistémico a cuestiones de hecho que de otro modo serían difíciles o imposibles de discernir” (Elgin, 2007, p. 39). Elgin señala que los modelos idealizados no dan acceso epistémico al mundo describiéndolo (pues, desde su punto de vista, lo describen de un modo no verídico), sino ejemplificando algunas de las características de los fenómenos que representan. La noción de “ejemplificación” que adopta Elgin tiene la siguiente definición técnica: una representación ejemplifica la propiedad F si y sólo si es un caso de esa propiedad y, además, hace referencia a la propiedad resaltándola (Cfr., Elgin, 2009, p. 81). Elgin presenta un ejemplo sencillo para ilustrar esta cuestión: las muestras de colores que presentan los catálogos comerciales de pinturas. Estos pedazos de cartón con un color impreso ejemplifican a las pinturas por las que están, porque son un caso de la propiedad pertinente (el color) y además resaltan esta propiedad al estar incluidas precisamente en un catálogo que tiene como finalidad presentar los colores de las pinturas.

Sin embargo, al considerar ejemplos más cercanos a la ciencia surgen dificultades a la hora de explicar cómo encaja esta noción de ejemplificación. Consideremos, por ejemplo, el caso del modelo predador-presa de Volterra. Según Elgin, los modelos idealizados proveen acceso epistémico a los fenómenos, si aquellos ejemplifican algunas propiedades del sistema que supuestamente representan. Dados los fines del modelo de Volterra, todo indica que éste debe proporcionar acceso cognitivo al menos a la interacción de las poblaciones de presas y predadores. Sin embargo, es dudoso que el modelo matemático mismo ejemplifique, en el sentido de Elgin, esa interacción; en todo caso parece que el modelo da acceso cognitivo a esta interacción mediante una descripción matemática. Quizás otro ejemplo deje más claro este punto. En la mecánica newtoniana la aceleración de un cuerpo está dada por la relación matemática de fuerza sobre masa. Es decir, la ecuación “ $a=F/m$ ” es un modelo idealizado matemático que representa la aceleración de los objetos.⁵³ De acuerdo con la propuesta de Elgin los modelos idealizados ejemplifican las propiedades de los objetos que representan. Esto significa que la fórmula de la aceleración debe ella misma tener la propiedad de la aceleración, lo cual resulta simplemente absurdo, pues los objetos matemáticos no pueden ejemplificar la propiedad de la aceleración. La cuestión sobre la que quiero hacer énfasis es que la ejemplificación de la que habla Elgin sólo parece tener sentido cuando tratamos con representaciones *materiales* (como el caso del catálogo) e incluso con modelos a escala o aumentados (por ejemplo, el modelo de doble hélice del ADN); sólo en estos casos parece que tiene sentido hablar de la ejemplificación de propiedades. Pero cuando nos enfrentamos a modelos *abstractos*, como los modelos matemáticos, resulta mucho más natural hablar de una *descripción* o de una *representación* de las propiedades del fenómeno u objeto que sea el objetivo del modelo. Parece que Elgin debería dar una explicación detallada de cómo

⁵³ Esta fórmula es un modelo idealizado en la medida que ignora muchos factores que también influyen en la aceleración de los objetos reales, por ejemplo la fricción que el ambiente ejerce sobre el cuerpo.

efectivamente la ejemplificación también es útil para explicar el acceso epistémico que proporcionan los modelos idealizados abstractos, en particular los modelos matemáticos.

3.4. En contra de la veracidad de las leyes científicas. La postura de Cartwright.

Hasta aquí he expuesto la respuesta que da Elgin a las cuestiones de en qué consiste el entendimiento científico y cuál es el logro epistémico asociado con el uso de modelos idealizados en la ciencia. También he identificado tres obstáculos que debería sortear la propuesta de Elgin para que podamos considerarla viable: 1) debe aceptar que la eliminación de la idealización en la ciencia a veces es un desiderátum de la actividad científica, como lo muestra el caso de van der Waals; 2) debe conciliar su definición de entendimiento con el desiderátum de que incluso los niños pequeños con una visión simple de un tema tienen al menos un entendimiento rudimentario del tema en cuestión y 3) debe explicar cómo incluso los modelos abstractos, como los modelos matemáticos, nos dan acceso epistémico a algunas propiedades del mundo mediante la ejemplificación.

Estos problemas, especialmente (3), bien pueden presentar serias dificultades para la propuesta de Elgin. Sin embargo, concedamos que podría elaborarse alguna sofisticación de la teoría de Elgin que los supere. En esta investigación no quiero profundizar sobre estos problemas. Quisiera enfocar mi crítica a esta postura en otra de sus características. Recordemos que la propuesta particular de Elgin sobre cuál es el logro epistémico asociado con los modelos idealizados se trajo a cuento para ilustrar un tipo de estrategia que se puede adoptar frente a este problema. Esta estrategia general consiste en aceptar que los modelos idealizados efectivamente son representaciones no-verídicas del mundo, lo cual conduce a los defensores de esta propuesta a identificar el logro cognitivo epistémico asociado con los modelos idealizados con un logro no fáctico. Me parece que es posible argumentar que la estrategia general (ejemplificada por Elgin) descansa en un supuesto erróneo que, en principio, pone en

cuestión la viabilidad de la estrategia en general. Me refiero al supuesto de que los modelos idealizados son representaciones no verídicas del mundo.

Este supuesto, que Elgin acepta sin someterlo a una consideración rigurosa, tiene su origen histórico en las consideraciones que Nancy Cartwright presenta en su libro *How the Laws of Physics Lie*. Ahí, ella arremete no sólo en contra de la idea de que la única meta de la actividad científica sea la descripción verdadera y completa del mundo, sino que también arremete en contra de la idea de que sea siquiera posible que nuestras mejores representaciones científicas del mundo sean verdaderas.

Antes de pasar a explicar con más detalle la postura de Cartwright me gustaría señalar una diferencia sutil entre dos tipos de errores representacionales: los errores y las distorsiones intencionales. Ambos son casos de representaciones no verídicas, es decir representaciones que presentan al mundo de un modo como éste no es. Sin embargo, un error es un caso de representación cuya incorrección no es producto de la intención del agente que crea la representación. Mientras que una distorsión intencional es una representación cuya incorrección se debe precisamente a la intención del agente, quien intencionalmente introduce algún tipo de distorsión en la representación con algún propósito, por ejemplo con la finalidad de volver más tratables los problemas desde un punto de vista computacional. Por ejemplo, una pintura de María Antonieta que ubica un lunar en un lugar incorrecto es un error, una representación errónea. Mientras que una caricatura de María Antonieta, que representa algunos de sus rasgos de una manera muy distorsionada, no es un error sino una distorsión intencional.⁵⁴

Es importante señalar que la tesis de Elgin y Cartwright que señala que la ciencia está llena de representaciones no verídicas debe interpretarse asumiendo que estas representaciones son

⁵⁴ Paul Teller (2009) presenta una distinción similar a ésta, sólo que Teller llama “ficciones” a lo que yo llamo distorsiones intencionales. Por lo demás, su distinción es consistente con lo que he dicho aquí. Él señala que “una ficción no es un error. Un error es una mala representación [*misrepresentation*] que ocurre sin querer, donde el autor de la representación no es consciente de la divergencia de la veracidad. Un discurso ficcional [...] también involucra una divergencia de una elucidación verídica, pero para calificar como ficcional, la divergencia debe de ser a propósito...” (Teller, 2009, p. 241).

distorsiones intencionales y no meros errores. Su idea es que los científicos introducen intencionalmente representaciones no verídicas.

Ahora bien, la postura de Cartwright surge a partir de un interés en la noción de explicación científica. Sería un error creer que el objetivo principal de Cartwright se reduce a elaborar alguna crítica en contra de una visión veritista de la ciencia, es decir una visión según la cual la verdad es la meta epistémica fundamental de la actividad científica. Su principal motivación consiste en argumentar a favor de un modelo causal de explicación y rechazar, en particular, el modelo basado en leyes de Hempel, es decir el modelo nomológico-deductivo e inductivo-estadístico. De acuerdo con Cartwright la explicación adecuada de un fenómeno consiste en identificar su causa. Por ejemplo, sólo damos una explicación adecuada de por qué murió mi limonero si apelamos a la causa efectiva de su muerte, por ejemplo al hecho de que el agua se acumuló en la base de la planta, de un modo tal que esta acumulación de agua es la causa de su muerte (*Cfr.*, Cartwright, 1983b, p. 91).⁵⁵

Hay que recordar que el modelo de explicación nomológico-deductivo tiene la característica de exigir que las premisas del argumento explicativo sean verdaderas.⁵⁶ Cartwright argumenta que precisamente este es uno de los puntos más débiles de este modelo explicativo pues, argumenta, las leyes científicas (que deben figurar necesariamente como premisas en una explicación nomológica deductiva) típicamente no son verdaderas. Y peor aún, si fuera posible elaborar leyes científicas verdaderas, su poder explicativo sería muy pobre. Cartwright identifica a “la postura fáctica de las leyes” como aquella que sostiene que las leyes de la naturaleza que la ciencia descubre y utiliza describen hechos acerca de la realidad (*Cfr.*

⁵⁵ Cartwright sostiene que la adopción de un modelo de explicación causal tiene ciertas ventajas sobre la adopción de un modelo nomológico-deductivo. Entre otras, la adopción del modelo causal nos permite inferir la verdad del efecto a partir de la causa, mientras que el modelo N-D no nos permite inferir la verdad de las leyes involucradas en una explicación, incluso cuando la identifiquemos como la mejor explicación. Esta y otras líneas de razonamiento a favor de su postura se desarrollan con detalle en su “When Inference Leads to Explanation” (1983c).

⁵⁶ Aunque propuse una modificación en la pp. 42-43 que relajaba esta exigencia y pedía meramente que las premisas estuvieran epistémicamente justificadas.

Cartwright, 1983b, p. 54). Ésta postura será el objetivo de su crítica y ofrecerá razones para rechazarla. El problema que ella encuentra con aceptar esta postura es que nos lleva al siguiente dilema, que ya presenté superficialmente en el párrafo anterior: “interpretadas como descripciones de hechos, [las leyes fundamentales de la física] son falsas; enmendadas para ser verdaderas, pierden su fuerza explicativa fundamental” (Cartwright, 1983b, p. 54). Es importante señalar que en esta afirmación Cartwright insinúa que tanto la verdad como la explicación son valores importantes para la ciencia y, más aún, que la ciencia está condenada a no conseguir ambas al mismo tiempo: entre más nos acercamos a alcanzar la meta de la verdad, más nos alejamos del poder explicativo y entre más nos acercamos al poder explicativo, más nos alejamos de la verdad.

Podemos ilustrar esta tesis tomando en cuenta el ejemplo paradigmático que presenta Cartwright: la ley de la gravitación universal. Esta ley dice, según la formulación estándar, que “dos cuerpos ejercen una fuerza entre ellos que varía directamente con el producto de sus masas y que varía inversamente con el cuadrado de la distancia que los separa” (Cartwright, 1983b, p. 57). El problema con la veracidad de esta ley consiste en que, estrictamente, no describe de manera la manera en la que se atraen los cuerpos, pues también la carga eléctrica tiene influencia en la manera en que los cuerpos se atraen, en particular la ley de Coulomb sostiene que la electricidad ejerce también una fuerza, que es inversamente proporcional al cuadrado de la distancia entre las cargas de los cuerpos (*Cfr.*, Cartwright, Ídem). Esto significa que la fuerza de atracción entre dos cuerpos nunca será descrita con precisión ni por la ley de la gravitación ni por la ley de Coulomb. Lo cual significa, a su vez, que ninguna de las dos leyes es universalmente aplicable: la ley de la gravitación describe de manera precisa la atracción entre dos cuerpos cuya atracción a causa de sus carga eléctricas es insignificante (por ejemplo, la atracción entre el sol y la tierra), mientras que la ley de Coulomb describirá correctamente la atracción entre los electrones y protones de un átomo, donde la fuerza gravitacional resulta

poco relevante. Así, hay muchos casos en los que tanto la ley de Coulomb como la ley de la gravitación dan una descripción incorrecta de la manera en que se atraen dos cuerpos.

Sin embargo, es posible proporcionar una versión modificada de estas leyes que sí son verdaderas. Basta con que añadamos un modificador *ceteris paribus* al inicio de las leyes que dijera algo así como “en ausencia de otras fuerzas además de la gravitacional (o de la eléctrica), dos cuerpos se atraen con una fuerza...”. La propia Cartwright acepta que las leyes en cuestión modificadas de esta forma efectivamente son verdaderas (Cartwright, 1983b, p. 57). Pero el problema que Cartwright identifica en ellas consiste en que este tipo de leyes tienen un poder explicativo muy limitado. Dado que, por ejemplo, la ley de la gravitación modificada nos dice cuál es la fuerza de atracción entre dos cuerpos sólo en las situaciones en las que no hay otras fuerzas presentes, entonces únicamente podemos utilizarla para explicar fenómenos en los que se cumple la cláusula *ceteris paribus*. Sin embargo, de acuerdo con Cartwright, esto nunca sucederá: por mínima que sea la influencia de la fuerza eléctrica, es casi seguro que siempre estará ahí, de modo que la cláusula *ceteris paribus* nunca se satisfará. Puesto que la ley modificada con la cláusula *ceteris paribus* describe condiciones ideales, su aplicación será limitadísima o quizás incluso nula.

La postura fáctica de las leyes enfrenta, por lo tanto, un problema: o bien interpreta a las leyes de la naturaleza de un modo literal, lo cual permite utilizarlas para explicar una gran gama de fenómenos o bien las interpreta de un modo no literal, añadiendo una cláusula *ceteris paribus*, manteniendo así su verdad, pero restringiendo su alcance explicativo. De manera simple: si modificamos las leyes para que sean verdaderas, explicaremos poco y si las dejamos como están, las leyes serán falsas pero explicarán mucho.

Expongo todo esto de una manera muy superficial porque el interés central en este trabajo no hace necesario que revise con detalle la dialéctica entre Cartwright y el defensor de la postura fáctica de las leyes. Para los propósitos de esta investigación basta con que quede claro

que Cartwright argumenta en contra de la postura fáctica de las leyes basándose en la tesis de que existe una incompatibilidad entre la verdad de las leyes de la naturaleza y su poder explicativo.

Elgin, por su parte, también sostiene que muchas de las representaciones que utiliza la ciencia son no verídicas. Y esgrime argumentos similares a los de Cartwright:

La ciencia se distancia a sí misma aún más de los fenómenos cuando recurre a modelos, idealizaciones y experimentos mentales. Los modelos son representaciones esquemáticas que resaltan características significantes, al tiempo que prescinden de complicaciones irrelevantes. [...] Pueden ser radicalmente incompletos, representando sólo aspectos seleccionados de los fenómenos. Estrictamente y literalmente, no describen nada en el mundo (Elgin, 2006, p. 209).

De este modo, tanto Elgin como Cartwright argumentan que muchas de las representaciones que se utilizan normalmente en la ciencia son no verídicas, basándose en el hecho de que las representaciones son incompletas (en el sentido de que ignoran algunos factores de los fenómenos reales), lo cual significa que *estrictamente* no representan ningún fenómeno real. Por ejemplo, la ley de los gases ideales supone que las partículas de los gases no ejercen entre sí ninguna fuerza de atracción. Esta fuerza de hecho está presente en los fenómenos reales, simplemente sucede que la ley ignora tal hecho por considerarlo irrelevante. Esto significa, argumenta Elgin, que la ley de los gases ideales estrictamente no representa ningún gas real, pues ignora la atracción de las moléculas que lo conforman.

Ahora bien, parece claro que el veredicto por parte de Elgin y Cartwright de que este tipo de representaciones científicas son estrictamente falsas depende de su adopción de una teoría de la representación muy estricta. A continuación sostendré que no hay buenas razones para creer que efectivamente uno debe aceptar esta teoría de la representación. Argumentaré que hay

alternativas más permisivas y que parecen hacer mayor justicia a la actividad científica, en el sentido de que la describe mejor. La existencia de esta teoría alternativa de la representación permite argumentar que la estrategia general que ejemplifica la postura particular de Elgin descansa en un supuesto incorrecto, a saber que los modelos idealizados son representaciones no verídicas del mundo.

Es importante señalar que si logro establecer que esto último es correcto, entonces el camino quedará libre para la adopción de una explicación del tipo (b) sobre el logro epistémico asociado con los modelos idealizados. Esta estrategia, hay que recordar, consiste en rechazar que los modelos idealizados sean representaciones no verídicas y, por lo tanto, permite identificar con el conocimiento o con otro logro epistémico fáctico al logro epistémico asociado con ellos.

3. 5. Una teoría alternativa de la representación

He explicado cómo la postura de Elgin en particular y la estrategia general que ejemplifica depende crucialmente de que una buena parte de las representaciones científicas sean no-verídicas, en el sentido de que sean distorsiones intencionales y no meros errores. A continuación sostendré que una teoría de la representación más permisiva tiene como consecuencia que el tipo de representaciones supuestamente no verídicas que presentan Elgin y Cartwright pueden de hecho entenderse como representaciones verídicas y, además, sostendré que adoptar una teoría de ese tipo hace mayor justicia a la práctica científica. Finalmente, argumentaré que esa teoría de la representación proporciona una explicación aceptable de por qué los modelos idealizados arrojan predicciones y descripciones correctas.

No es una tarea fácil identificar de manera precisa qué teoría de la representación subyace a las posturas de Elgin y Cartwright, pues nunca la presentan de manera explícita. Sin embargo, a partir de sus razonamientos, es posible inferir que asumen una teoría según la cual una

representación no puede ser verídica si no representa de manera *precisa todas* las características del objeto o del fenómeno que busca representar. Así, de acuerdo con una postura como ésta, la ley de los gases ideales es falsa, por ejemplo, porque no toma en cuenta el hecho de que las moléculas del gas efectivamente tienen volumen y ejercen entre sí una fuerza de atracción. O, por ejemplo, según esta teoría de la representación, la ley de la gravitación universal es falsa porque no toma en cuenta la fuerza de atracción entre los cuerpos producto de su carga eléctrica. Estos ejemplos muestran que, de acuerdo con esta teoría, una representación sólo es verídica si tiene el grado máximo de *fidelidad* y de *completud* con respecto a los hechos. A continuación explicaré que es la fidelidad y la completud de una representación.⁵⁷

La *fidelidad* de una representación consiste en el grado de precisión con que ésta representa a los objetos o fenómenos. Por ejemplo, la fidelidad de la oración “Sebastián mide 185.3cm” tiene que ver con qué tan precisa es la representación, que aquí está directamente relacionada con cuál es la unidad de medida más pequeña que captura la representación, en este caso los milímetros. En un contexto cotidiano, por ejemplo en un examen médico, este grado de fidelidad resulta satisfactorio. Sin embargo, de acuerdo con la postura que suponen Cartwright y Elgin, cualquier representación debe tener el grado máximo de fidelidad, aunque el contexto en el que ésta se utiliza no exija tal grado de fidelidad. De acuerdo con ellas, por ejemplo, una representación lingüística verídica sobre la altura de Sebastián debería tener una fidelidad que se vea reflejada, quizás, hasta el grado de los micrómetros o los nanómetros. Por otro lado, la *completud* de una representación tiene que ver con cuántas características del objeto o fenómeno son representadas. Por ejemplo, el modelo de Volterra era, en este sentido, incompleto, pues dejaba fuera de la representación muchas características que sí estaban presentes en el fenómeno real: la densidad de la población, las relaciones espaciales y la

⁵⁷ Sobre la distinción entre completud y fidelidad en una representación véase “Fiction, Fictionalization and Truth in Science” (2009), especialmente la sección 2 y “Representation in Science” (2008), ambos de Paul Teller.

interacción entre otras especies no incluidas dentro del grupo de presas o predadores, entre otras. De acuerdo con la teoría extrema de la representación que sostienen Elgin y Cartwright una representación sólo es verdadera si, además de representar con máxima precisión las características del objeto o fenómeno a representar, representa todas y cada una de las características del objeto o fenómeno.

Aquí quiero argumentar a favor de una teoría de la representación distinta, que toma en cuenta el contexto en el que se usa una representación para determinar si la representación es verídica o no. La llamaremos *teoría contextual de la representación*. Una de las consecuencias más importantes de esta teoría es que una representación puede ser verídica aun cuando no sea completamente fiel ni absolutamente completa. van Fraassen, por ejemplo, argumenta a favor de una postura como ésta.⁵⁸ Él sostiene que “una representación está hecha con un propósito o meta en mente y está gobernada por criterios de adecuación que tienen que ver con esa meta, que guía sus medios y selectividad” (van Fraassen, 2008, p. 7). Es decir, el criterio para determinar si una representación es verídica está ligado con la meta o el propósito con el que esa representación se construye. La meta de la representación determina, al menos parcialmente, cuál es el medio que se usará para llevar a cabo la representación, cuál será el grado de fidelidad que se busca y también qué características del fenómeno u objeto se representarán. En pocas palabras, el contexto en el que se crea la representación y el modo en que la representación se utiliza determina qué criterio hay que adoptar para juzgar la veracidad de la representación.

Quizás un ejemplo de la vida cotidiana sea útil para aclarar cómo los criterios de adecuación de las representaciones suelen ser relativos a contextos y metas: consideremos el

⁵⁸ Argumenta de una manera especialmente enérgica a favor de una teoría de la representación de este tipo en su libro *Scientific Representation: Paradoxes of Perspective*. Otros filósofos han defendido teorías similares. Véase especialmente *Science and Partial Truth. A Unitary Approach to Models and Scientific Reasoning* (2003) de Newton C. A. Da Costa y Steven French, así como “Fictions, Fictionalization and Truth in Science” (2009) y *Representation in Science* (2008) de Paul Teller.

caso de un mapa de la red del metro del Distrito Federal. Para el uso normal de este tipo de mapas basta con que se represente correctamente el orden de las estaciones en la línea correspondiente y que las intersecciones entre las líneas se ubiquen en donde de hecho están. Basta con que el mapa represente correctamente estas características para considerarlo una representación verídica de la red del metro del Distrito Federal, dado el contexto en que se produce y se utiliza la representación. Es posible que el mapa no represente de manera correcta la distancia entre las estaciones ni el tamaño de las estaciones mismas, sin embargo, esto no contará necesariamente como un error en la representación. De hecho, dados los intereses asociados con este tipo de mapa, representar de manera parcial las características del sistema del metro parece ser la opción óptima, pues incluir más información podría resultar en un mapa tan abigarrado que no se pueda comprender fácilmente en dónde está cada estación.

Esta teoría va incluso más lejos. Sostiene que, en ocasiones, la representación *óptima* exige que se distorsione intencionalmente la representación. van Fraassen, por ejemplo, señala que “la distorsión, la infidelidad y la falta de similitud en algún aspecto, puede en general ser crucial para el éxito de una representación. Esto no significa que no se requiera similitud en algún otro aspecto” (van Fraassen, 2008, p. 13). De hecho, lo más común es que los mapas del metro representen a las estaciones de un modo distorsionado: regularmente las estaciones se representan mucho más grandes de lo que son en realidad, pero esto se hace para resaltar el hecho de que en un punto particular hay una estación. Así, en el caso del mapa del metro, para tener una representación adecuada, se requiere que se representen fielmente algunas características (por ejemplo el orden de las estaciones) y que se distorsionen otras (como el tamaño de las estaciones). En pocas palabras, estas consideraciones muestran que en ocasiones es necesario introducir distorsiones y omitir elementos en una representación para que ésta represente correctamente su objeto. Así, esta teoría de la representación sostiene que las

representaciones incompletas y las que no son completamente fieles también pueden ser verídicas.

En este punto uno podría preguntarse cuál es la diferencia entre la postura de Elgin y la que presento aquí, pues después de todo, de acuerdo con ambas posturas este tipo de representaciones permiten un acceso cognitivo a los hechos. Hay, sin embargo, dos diferencias cruciales entre estas posturas. En primer lugar, Elgin sostiene que los modelos idealizados nos dan acceso cognitivo a los hechos mediante la ejemplificación, mientras que la teoría alternativa que defiende aquí nos permite adoptar la postura de que los modelos idealizados nos dan acceso cognitivo a los hechos mediante la representación. Ya señalé más arriba (en las páginas 102-103) que la noción de ejemplificación que defiende Elgin tiene un problema, en particular que no parece explicar correctamente cómo funcionan las representaciones matemáticas. La segunda diferencia consiste en que la teoría de la representación que subyace a la teoría de Elgin le impide considerar que los modelos idealizados sean representaciones verídicas del mundo, mientras que la teoría que yo asumo sí me permite considerarlas como verídicas. Por esta razón Elgin se ve obligada a adoptar una interpretación del logro epistémico asociado con uso de estos modelos como si fuera un logro no fáctico. La adopción de la teoría contextual de la representación me permite interpretar a este logro epistémico como si fuera fáctico. En la sección (3.2) presenté dos posturas sobre cuál es el logro epistémico asociado con el uso de modelos idealizados que son consistentes con el principio de que no haya premisas falsas.

Continuando con la discusión de la teoría contextual de la representación, es pertinente preguntarse ¿Cómo puede una representación ser verídica si no representa todos los aspectos de su objeto ni los representa con la mayor fidelidad posible? La respuesta más sencilla consiste en

señalar que la representación regularmente es *selectiva* y que esto no implica su no veracidad.⁵⁹ Una representación selectiva representa sólo algunos aspectos del objeto y lo hace de una manera que no es completamente fiel. Desde el punto de vista de una teoría de la representación contextual, esto no es una razón suficiente para que una representación no pueda ser verídica. La veracidad de una representación selectiva depende de la representación más o menos fiel de *sólo algunas* de las características de un objeto. Ahora bien, qué características son las que deben representarse y cuál es el grado de fidelidad que se exige a la representación para que sea verídica es una cuestión que varía con el contexto y con la meta o finalidad de la representación. Si aceptamos la tesis general de que la representación regularmente es parcial (es decir, imprecisa e incompleta), entonces el mero hecho de que alguna representación sea imprecisa o incompleta no implica que la representación es no verídica, pues puede ser verídica dada la finalidad que se persigue con ella en algún contexto.

Esta teoría de la representación sostiene que la parcialidad de una representación está en función de los intereses con los que se usa esa representación. Retomando el ejemplo del mapa del metro, tenemos que nuestros intereses determinan qué características deben representarse correctamente (el orden de las estaciones y la ubicación de las intersecciones) y qué grado de precisión exigiremos para la representación. En la medida que un mapa cumpla con estas exigencias diremos que es una representación verídica. Es cierto que hay un tipo de corrección que establece que una representación es verídica sólo si representa con el máximo grado de fidelidad y completud a su objeto. Podemos llamar a este tipo de veracidad “verdad completa”. Podemos, por otro lado, llamar “verdad parcial” a la noción de verdad que surge de una teoría contextual de la representación. Sin embargo, quiero argumentar que la verdad completa es sólo un tipo de veracidad para las representaciones y, de manera más importante, que en la actividad

⁵⁹ Paul Teller (2008) argumenta que la selectividad y la inexactitud en las representaciones está presente en todos los aspectos de la actividad humana. Él parece creer que ésta es una razón suficiente para favorecer una interpretación de este tipo de representaciones que permita que sean verídicas.

científica raramente se asume como criterio de veracidad el criterio asociado con la verdad parcial y no con la verdad completa.

A continuación explicaré brevemente cómo se evalúa la veracidad de una representación de acuerdo con una teoría contextual de la representación. Dados ciertos intereses, uno se compromete a representar correctamente un conjunto de características del objeto o fenómeno en cuestión y se establece cuál es el grado de precisión (o fineza de grano) que se exigirá a la representación. Una vez especificado esto, se evalúa a la representación como verídica o no verídica sólo en la medida en que represente las características que se ha comprometido a representar correctamente y que las represente con el grado de fidelidad con el que se ha comprometido. Por ejemplo, consideremos una representación material a escala de un edificio. Supongamos que, dados los intereses del contexto, se adoptan los compromisos de representar correctamente sólo su estructura (pero no sus materiales, su color ni su decoración) y con una precisión que llega sólo a los centímetros con la finalidad de presentar el modelo como una maqueta de un proyecto de construcción. Sólo en la medida en que se satisfagan estos compromisos representacionales diremos que la representación es verídica. El mero hecho de que haya un conjunto de características que no se incluyen en la representación o que no son representadas con completa precisión no habla en contra de la veracidad de la representación misma, sino sólo en contra de su completud y de su fidelidad: una representación selectiva sólo describe correctamente un conjunto limitado de las características de un objeto y las describe con una precisión limitada y basta esto para que podamos hablar de representaciones selectivas verídicas.⁶⁰

⁶⁰ Da Costa y French dan una presentación formal de la noción de verdad selectiva en su libro *Science and Partial Truth. A Unitary Approach to Models and Scientific Reasoning* (2003). Puede encontrarse una presentación aún más precisa de esta noción en el artículo de Mikenberg, Da Costa y Chuaqui “Pragmatic Truth and Approximation to Truth” (1986).

Quisiera ahora sugerir que una manera de presentar esta teoría de la representación consiste en sostener la tesis de que el predicado “es verídico” está indexado a un contexto de uso. De acuerdo con la filosofía del lenguaje contemporánea hay una amplia variedad de términos cuyo referente varía en función del contexto en que se utiliza el término. Estos términos son conocidos como *deícticos*. La característica esencial de los deícticos consiste precisamente en su sensibilidad al contexto: “diferentes aseveraciones del mismo deíctico designan cosas distintas, porque lo que se designa depende no sólo del significado asociado con la expresión, sino también de hechos sobre la aseveración” (Perry, 1998, p 586). De acuerdo con esta postura, el contexto asociado con la aseveración de un término indexical determina parcialmente cuál es el referente del término. El contexto consiste, como bien señala Perry, en algunos hechos sobre la aseveración misma, como el lugar y la hora en que se lleva a cabo la aseveración, quién es el sujeto que la emite y quién aquél al que se dirige, etc. Es muy importante señalar que, del mismo modo en que la referencia de un deíctico varía de contexto a contexto, también el valor de verdad⁶¹ de una oración que hace uso de un deíctico variará de contexto a contexto: “las oraciones que contienen deícticos pueden usarse para decir cosas diferentes en ocasiones diferentes” (Perry, 1998, p. 586). Ejemplos paradigmáticos de este tipo de términos son “yo”, “tú”, “mi”, “suyo”, “aquí”, “ahora”, “éste”, “ese”, “hoy”, entre otros. Quizás sea importante señalar que si bien todos estos términos son deícticos hay una distinción importante entre ellos. Los demostrativos son los deícticos que requieren de una demostración asociada para determinar su referente en todos los contextos de uso. Un término paradigmático de este tipo es “eso”. Por otro lado, los deícticos puros son aquellos deícticos que no requieren de una demostración asociada para determinar su referente, las reglas lingüísticas que gobiernan su uso bastan para determinar cuál es su referente en cada contexto. Para los

⁶¹ En el caso de las oraciones, su referencia es su valor de verdad. Así, resulta que la referencia de los deícticos (los objetos a los que refieren) y la de las oraciones que hacen uso de ellos (sus valores de verdad) varían de contexto a contexto.

deícticos puros, cualquier demostración asociada con ellos está de más. Un término de este tipo es “Yo”. (Cfr, Kaplan, 1989, pp. 490-491).

De acuerdo con la teoría clásica de Kaplan, los deícticos se caracterizan por tener dos tipos de significado: contenido y carácter. El contenido de un deíctico es *lo que se dice* con él. Del mismo modo, el contenido de una aseveración que hace uso de un deíctico es lo que se dice con esa aseveración. El contenido de una aseveración es una proposición y ésta se evalúa en las circunstancias de evaluación para arrojar un valor de verdad para la proposición. Por ejemplo si yo digo hoy “fui al centro ayer” y tú aseveras la misma oración mañana, estaremos diciendo cosas diferentes. Por lo tanto, el contenido de nuestras aseveraciones será distinto, es decir las proposiciones expresadas por nuestras aseveraciones serán distintas. El carácter, por otro lado, es el tipo de significado que determina el contenido de una aseveración en distintos contextos. El carácter de un deíctico se determina mediante convenciones lingüísticas y, en última instancia, determina el contenido de la expresión en cada contexto. Por ejemplo, la regla

“Yo” refiere al hablante que lo usa

es una regla que determina el carácter de esta expresión (“yo”) y, en última instancia determina cuál es su contenido en cada contexto.⁶²

La teoría de la representación que quiero defender aquí trata al predicado “es verídico” como si fuera un deíctico, esto es como un predicado cuyo referente varía de contexto a contexto. En la breve presentación que he hecho aquí de la teoría de los deícticos sólo he hablado de deícticos que refieren a objetos y oraciones en las que esos deícticos figuran, pero no he dicho nada sobre los predicados. Hay también posturas en filosofía del lenguaje que

⁶² Esta caracterización del significado de los deícticos según la cual consiste en contenido y carácter es controvertido. Por ejemplo, John Perry sostiene que la caracterización de “contenido” de Kaplan es bastante cuestionable y, propone, en su lugar dar una caracterización según la cual los deícticos tienen tres niveles de contenido. Véase Perry (1998). En esta investigación no es necesario entrar a discutir con detalle estas cuestiones.

sostienen que algunos predicados son deícticos, por ejemplo “adinerado”, “alto”, “extraordinario”. Daniel Rothschild y Gabriel Segal (2009) presentan una semántica para predicados deícticos, en particular para predicados de color. Aquí no la presentaré detalladamente, sino sólo superficialmente. Sugiero que el predicado “es verídico” puede funcionar del modo en que funcionan los predicados deícticos que caracterizan Rothschild y Segal. Si se aplica su propuesta a este caso particular, tendríamos lo siguiente. Habría que agrupar en una clase todos los casos de una expresión que ocurre en un contexto. Agrupamos, por ejemplo, en un tipo sintáctico las ocurrencias de “es verídico” que ocurren cuando lo usamos en el contexto del mapa de la línea del metro. Ningún caso de este tipo sintáctico puede ocurrir en otro contexto. Se indica el tipo sintáctico añadiendo índices. Así, todos los casos de “es verídico” que ocurren en el mismo contexto reciben el mismo índice, por ejemplo “es verídico_n” y nos referimos al contexto en cuestión como “contexto-n”. La extensión que recibe cada tipo sintáctico que está ligado a un contexto de “es verídico” es relativa a su contexto: una representación satisface “es verídico_n” en el contexto-n si cuenta como verídico dados los estándares del contexto. (*Cfr.* Rothschild y Segal, 2009). Así, para cada contexto de uso del predicado “es verídico”, su referencia estará fijada según los estándares del contexto.⁶³ En este caso los estándares del contexto son los compromisos representacionales, es decir la especificación de cuáles aspectos del fenómeno se van a representar y con qué grado de fidelidad se representarán. Cuáles son los estándares particulares dependerá de cada contexto particular.

Hay por lo tanto, un marco teórico bien establecido en la filosofía del lenguaje que puede dar mucha mayor sustancia y precisión a la teoría contextual de la representación que definiendo aquí. En esta ocasión no busco dar una elucidación completa y precisa de la manera en que

⁶³ Véase Rothschild y Segal (2009) para una presentación detallada de la semántica de los predicados deícticos.

funcionaría este predicado si es entendido como un deíctico. Simplemente deseo sugerir que, a primera vista, hay una manera en que podría explicarse en estos términos.

Después de la presentación de esta sugerencia, continuaré con el análisis de la teoría contextualista de la representación. A continuación argumentaré que esta teoría efectivamente explica de una manera aceptable la manera en que se usan los modelos idealizados en la actividad científica. Es necesario preguntarnos en este punto si hay razones para creer que esta teoría describe correctamente la manera en la que se evalúa la veracidad de los modelos idealizados en los contextos científicos. Michael Weisberg da razones para creer que efectivamente esta teoría de la representación es la más adecuada para caracterizar estas representaciones científicas. Weisberg ofrece un análisis de en qué consisten los modelos científicos en su “Who is a Modeler?” (2007a). Ahí presenta una postura según la cual las intenciones relevantes del creador del modelo con respecto a cuál es la relación adecuada entre la representación (el modelo) y el mundo está dada por una *interpretación* del modelo. La interpretación, a su vez, está compuesta por cuatro elementos: una asignación, un alcance y dos tipos de criterios de fidelidad. La asignación consiste en hacer explícito cuál es el fenómeno o el objeto que se busca representar y, en particular, establece qué partes del modelo corresponden a qué partes del objeto. Y también se especifica qué partes del modelo no representan nada del objeto. El alcance del modelo establece qué características del objeto pretende representar el modelo. Finalmente, los criterios de fidelidad nos dicen qué tan similar al objeto debe ser un modelo para que se lo considere una representación verídica. Hay dos tipos de criterios de fidelidad: dinámicos y representacionales. La *fidelidad dinámica* nos dice qué tan cerca debe estar el resultado que nos da el modelo del resultado real del fenómeno para que lo aceptemos como un modelo adecuado. Por ejemplo, establecemos un margen de error para las predicciones que se obtienen a partir de la ley de los gases ideales. Si el comportamiento de los gases se desvía demasiado de la predicción, decimos que el modelo no

es adecuado. Por otro lado, la *fidelidad representacional* nos permite evaluar si el modelo hace las predicciones correctas por las razones correctas (Cfr. Weisberg, 2007a, pp. 218-221). Estos criterios de fidelidad especifican qué tan fielmente debe representar el modelo el proceso causal del fenómeno representado. Un ejemplo ayudará a clarificar a qué se refiere la fidelidad representacional. Supongamos que el modelo del sistema solar heliocéntrico y el modelo geocéntrico dan predicciones igual de exitosas. En particular imaginemos que dan predicciones exitosas del movimiento retrógrado de Marte. Sin embargo, el sistema heliocéntrico da una predicción exitosa por buenas razones, mientras que el sistema geocéntrico lo hace por malas razones. En otras palabras, el proceso causal del movimiento retrógrado de Marte es representado de un modo correcto por el sistema heliocéntrico, pero no por el sistema geocéntrico. Así, la fidelidad representacional del sistema heliocéntrico es mayor que la del sistema geocéntrico.

Hay que señalar que estas consideraciones surgen de una discusión teórica sobre la naturaleza de los modelos científicos. Sería bueno mostrar qué ocurre en la actividad científica misma. Para ilustrar cómo efectivamente los científicos son conscientes de la parcialidad de sus representaciones basta con echar un vistazo a cómo presenta la ley de los gases ideales un manual universitario introductorio de Química:

La ley de los gases ideales involucra todas las características importantes de un gas: su presión, su volumen, su número de moles y su temperatura. [...] Es importante reconocer que la ley de los gases ideales se basa en mediciones experimentales de las propiedades de los gases. De un gas que obedece esta ecuación se dice que se comporta *idealmente*. Esto es, esta ecuación define el comportamiento de un gas ideal. La mayoría de los gases obedecen esta ecuación cercanamente a presiones de aproximadamente 1 atm o menos, cuando la temperatura es aproximadamente 0 °C o mayor (Zumdahl, 2010, p. 419).

En este caso podemos ver cómo el autor dice de manera explícita cuáles son las condiciones en las que el modelo es aplicable (a saber, a presiones de 1 atm o menores y a temperaturas de 0°C o mayores) y también nos dice cuál es su alcance, es decir cuáles son las características que el modelo describe apropiadamente: la presión, el volumen, la temperatura y el número de moles del gas. Tenemos aquí un caso en el que se acepta explícitamente que el modelo arroja resultados precisos sólo en ciertas condiciones, y se acepta también que se representan adecuadamente sólo algunas características presentes en el fenómeno.

Estos dos casos nos muestran que en la discusión filosófica y en la actividad científica las representaciones idealizadas están presentes de manera generalizada. Sin embargo por sí mismo esto no implica que la teoría contextual de la representación sea correcta. Sin embargo sí argumentaré que esta teoría da sentido a la práctica científica, tal y como la presentan Weisberg y Zumdahl. Hay que recordar que en la sección 3.1, en este mismo capítulo, presenté brevemente dos modelos científicos: el modelo de Volterra y la ley de los gases ideales. Al presentarlos, logré identificar dos logros epistémicos asociados con el uso de esos modelos: la predicción exitosa y la descripción correcta. La pregunta que debemos responder es cómo se alcanzan estos logros epistémicos utilizando modelos idealizados como los que presentan Weisberg y Zumdahl. La explicación más directa y simple de este fenómeno consiste en señalar que los modelos idealizados, a pesar de ser representaciones parciales, representan al mundo de una manera verídica. De modo que cuando razonamos haciendo uso de estos modelos, arribamos a descripciones verídicas y a predicciones exitosas. De otro modo sería un misterio cómo se alcanzan estos logros epistémicos. Si los modelos idealizados fueran representaciones no-verídicas del mundo resultaría simplemente un misterio cómo la actividad científica, tomando como insumos representaciones no verídicas, arroja verdades. Así, la teoría contextual de la representación nos permite dar una explicación adecuada de estos aspectos de la actividad

científica. Por esta razón sostengo que esta teoría nos permite elaborar una imagen realista de la actividad científica.

Ahora bien, también es posible dar una explicación más detallada de por qué son típicamente aceptables en la actividad científica las representaciones parciales. Como señalé al inicio del capítulo 1, la actividad científica tiene una multitud de metas, entre las cuales se encuentran, al menos, la representación verídica y la explicación científica, pero también metas económicas, sociales y políticas, entre otras. Esto sugiere que sería un error pensar que los científicos tienen como única meta elaborar representaciones completísimas y muy fieles del mundo. Como muestra el ejemplo de Zumdahl, los científicos son perfectamente conscientes de las limitaciones de sus herramientas representacionales y precisamente por eso acuden a las idealizaciones: se dan cuenta de que, por diversas de razones, no pueden elaborar la representación más completa y más precisa de un fenómeno. Pero esto no significa que no busquen una representación que sea verídica, aunque sea parcial. Esto, en última instancia muestra de manera clara que la veracidad no es la única meta de la ciencia, pues regularmente se sacrifica la veracidad a favor de alguna otra meta científica. Quizás el ejemplo más claro es el de la capacidad computacional. Es cierto que la ciencia tiene como ideal elaborar representaciones más o menos precisas y completas de los fenómenos que estudia (*Cfr.* Weisberg, 2007b pp. 649-650), sin embargo, también debe tomar en cuenta si un modelo completo y preciso será computable. Así, las limitaciones computacionales regularmente restringen qué tipo de modelo se puede elaborar.⁶⁴ De igual modo, consideraciones de

⁶⁴ Winsberg (2009) presenta un ejemplo en el que claramente las capacidades computacionales restringen el tipo de modelo que se puede ofrecer. Él presenta el ejemplo un modelo “multi-escala” que se usa para estudiar las propiedades mecánicas de trozos sólidos de silicón. Este modelo incorpora tres niveles de descripción: mecánica cuántica (MC), dinámica molecular (DM) y mecánica de medios continuos (MMC). Es necesario integrar estos tres niveles de descripción porque cada uno es preciso y computable a distintas escalas. La MMC es la más computable, pero da resultados imprecisos cuando se aplica a escalas menores de una micra, pues asume que los sólidos son homogéneos incluso a estas escalas, lo cual simplemente es falso y resulta en predicciones equivocadas. El marco ideal para describir sólidos de estas dimensiones es la DM, pero cuestiones computacionales sólo permiten hacer cálculos entre 10^7 - 10^8 moléculas. En dimensiones lineales esto es sólo

simplicidad, de interés en los factores causales en un fenómeno o de interés en el poder explicativo nos pueden llevar a optar por adoptar representaciones incompletas o imprecisas, pero que ayuden a alcanzar específicamente alguna de estas metas. La adopción de modelos parciales no es en sí mismo una desventaja, pues, de nuevo, sólo si adoptamos una teoría de la representación muy estricta nos vemos obligados a aceptar que las representaciones parciales son necesariamente no verídicas.

Resumiendo, la teoría contextual de la representación es mucho más flexible que la que asumen Elgin y Cartwright. A su favor habla no sólo el hecho de que muchas prácticas humanas representacionales cotidianas son explicadas adecuadamente mediante esta teoría, sino también el hecho de que puede usarse para describir correctamente algunas partes de la actividad científica. Finalmente me gustaría señalar que una teoría contextual de la representación tiene la ventaja de dar una explicación sencilla y directa de cómo el uso de modelos idealizados tiene como resultado predicciones exitosas y descripciones verídicas. Como los modelos idealizados, desde el punto de vista de esta teoría, pueden representar verídicamente al mundo, por lo tanto también pueden describir verídicamente situaciones particulares y hacer predicciones exitosas. Por ejemplo, como la ley de los gases ideales describe verídicamente el comportamiento de los gases bajo ciertas circunstancias y para ciertos propósitos, una aplicación de esta ley a un caso particular puede tener como resultado descripciones verídicas y predicciones correctas, dadas ciertas circunstancias y ciertos

cerca de 50 nanómetros. Así DM es muy costosa computacionalmente y los métodos de MMC son demasiado imprecisos para sólidos de, por ejemplo, una micra de diámetro. Lo que se hace, entonces, es que se usa DM para modelar regiones donde hay grietas y se usa MMC para las regiones adyacentes, donde las cosas se mantienen suficientemente cerca del equilibrio. Pero hay otra complicación. Cuando las grietas se propagan por un sólido, el proceso involucra el rompimiento de enlaces químicos. De modo que los métodos de DM (que usan un modelo clásico de la interacción energética entre los átomos) no es fiable cerca del borde de la propagación de la grieta. Construir un modelo del rompimiento de los enlaces requiere de la MC. Pero aquí no se puede modelar más de 250 átomos a la vez, debido a cuestiones computacionales. Y así, se desarrolla un modelo que utiliza tres niveles de descripción que se aplican a diferentes regiones del sólido de silicón. Claramente, lo más deseable sería aplicar el modelo MC a todo el sólido. Sin embargo complicaciones computacionales llevan a incluir modelos distintos que son precisos sólo a otros niveles. Este ejemplo muestra de manera muy clara cómo las limitaciones computacionales restringen el modelo que se puede dar para representar un objeto.

propósitos. En última instancia, esta teoría puede explicar estos logros epistémicos porque caracteriza al proceso de utilizar modelos idealizados como uno que recibe como insumos representaciones verídicas y arroja, como consecuencia, más representaciones verídicas.

A continuación presentaré mi propuesta de en qué consiste el logro cognitivo asociado con el uso de modelos idealizados. Recordemos que la estrategia que adopto pertenece al tipo (b) de estrategias que sostiene que los modelos idealizados pueden ser representaciones verídicas y que identifican al logro epistémico asociado con éstos con el conocimiento o con otro logro epistémico fáctico. Ya he argumentado a favor de una teoría que permite aceptar que efectivamente los modelos idealizados pueden representar verídicamente al mundo. Ahora hace falta explicar con detalle en qué consiste el logro epistémico que estos proporcionan y por qué es posible interpretarlo fácticamente.

3.6. En qué consiste el logro epistémico asociado al uso de modelos idealizados

A continuación defenderé la postura de que el logro epistémico fundamental asociado con los modelos idealizados consiste simplemente en la obtención de conocimiento acerca del objeto que éstos representan. Claramente, no cualquier creencia que se obtenga haciendo uso de modelos idealizados constituirá conocimiento. Hay una variedad de condiciones extras que deben cumplirse: las creencias deben de haber sido formadas de manera adecuada, es decir de manera tal que no sean susceptibles de caer en casos Gettier, también deben de estar justificadas y ser verdaderas. El punto importante de la postura que presento es que, a diferencia de la postura de Elgin, permite que alcancemos conocimiento haciendo uso de modelos idealizados. Es claro que un científico descuidado que hace uso de modelos idealizados anticuados, que no lleva a cabo con cuidado sus experimentos y que razona de manera precipitada, difícilmente alcanzará conocimiento, a pesar de que alcance alguna creencia verdadera de vez en cuando. El que se obtenga conocimiento sobre alguna porción de

la realidad particular haciendo uso de modelos idealizados dependerá tanto del grado al que la ciencia ha avanzado así como de la calidad de la investigación científica que se lleva a cabo.

Diremos, por ejemplo, que la ley de los gases ideales proporciona conocimiento sobre el comportamiento de los gases (en particular sobre su presión, volumen temperatura y número de moles) bajo ciertas circunstancias de presión y densidad. El modelo predador-presa de Volterra nos da conocimiento de cómo interactúan poblaciones de predadores y presas en ciertas circunstancias; el mar Adriático es un caso particular en el que el modelo se aplica y, por lo tanto, su aplicación nos da conocimiento sobre la dinámica entre predadores y presas en el mar Adriático. Sin embargo, también puede darnos conocimiento sobre la dinámica de otros sistemas biológicos que se parezcan lo suficiente al mar adriático. Hay que aclarar qué significa decir, en este contexto, que un modelo idealizado proporciona creencias. Significa simplemente que un sujeto que conoce la manera en la que el modelo funciona puede utilizarlo para formar creencias a partir de un razonamiento que involucra la observación y la interpretación de los resultados de una aplicación del modelo a una situación particular. Por ejemplo, un científico puede aplicar la ley de los gases ideales a una tetera y determinar a qué presión sale el gas si la temperatura alcanza los 100°C. Este científico, mediante la aplicación e interpretación del modelo a esta situación particular puede formar la creencia de que el gas saldrá con una presión determinada.

Es muy importante notar que la diferencia entre mi postura y la de Elgin, que me permite adoptar la tesis de que los modelos idealizados nos dan conocimiento, es precisamente que mi postura adopta la tesis de que los modelos idealizados sí pueden ser representaciones verídicas de los fenómenos. De acuerdo con la discusión de la sección (3.2), si se acepta el principio de que no haya premisas falsas y, además, se sostiene que los modelos idealizados son representaciones no verídicas del mundo (tal y como hace Elgin), entonces *no se puede* adoptar la tesis de que los modelos idealizados nos dan conocimiento del mundo. Estos tres

compromisos teóricos simplemente son incompatibles. En cambio, una vez que se rechaza la idea de que los modelos idealizados son representación no verídicas, nada nos impide adoptar la tesis de que éstos pueden proporcionar conocimiento del mundo.

Debería resultar claro que sostener que el conocimiento es el logro asociado con el uso de modelos idealizados vuelve la tarea muy fácil para el teórico veritista, pues, como ya se señaló antes, el conocimiento es un logro epistémico fáctico: implica la verdad de lo que se conoce. El valor de este logro epistémico claramente es explicable bajo los términos de una teoría veritista: el conocimiento es valioso no porque tenga alguna relación con la creencia verdaderas, sino simplemente porque *está constituido* por creencias verdaderas.

Pero es claro que hay algunos modelos científicos que producen creencias falsas. Por ejemplo, el modelo planetario de Ptolomeo lleva a aceptar la creencia de que los planetas y el sol orbitan alrededor de la tierra, mientras que el modelo de Copérnico conduce a la creencia de que los planetas giran al rededor del sol en órbitas circulares, no elípticas. ¿Qué diría mi postura sobre estos casos? La respuesta más directa es que simplemente estas creencias no constituyen conocimiento. De nuevo, mi propuesta no está comprometida con la idea de que *siempre* que se usan modelos idealizados se alcanzará conocimiento. Además, es claro que si un modelo idealizado es falso, difícilmente podremos usarlo para conseguir conocimiento. Claramente, la teoría contextual de la representación que he adoptado no me compromete con la idea de que *todo* modelo idealizado sea verídico, sino simplemente con que *puede* serlo. Efectivamente, uno no cree que modelos idealizados descabellados nos lleven a adquirir conocimiento. Sin embargo, es claro que los modelos planetarios que acabo de presentar no son modelos descabellados. Resulta mucho más difícil dar una explicación de qué ocurre en estos casos. Es cierto que el modelo de Ptolomeo no representaba verídicamente la ubicación de la tierra en relación con los demás cuerpos del sistema solar, mientras que el modelo de Copérnico no representaba verídicamente la forma de las órbitas de los planetas. Sin embargo, algunas partes

de sus modelos representaban verídicamente al mundo. Por ejemplo, el modelo de Copérnico representaba de manera verídica el lugar de la tierra en el sistema solar. Es muy probable que esto ocurra con muchos otros modelos: que representen verídicamente algunas partes del fenómeno y no verídicamente otras partes, incluso desde el marco de la teoría contextualista de la representación. Al utilizar modelos idealizados (ya sean verídicos o no) uno puede arribar a creencias falsas. En muchos casos se podría argumentar que estas creencias falsas no constituyen ningún tipo de logro epistémico.⁶⁵ Sin embargo hay casos en los que parece que hay un logro epistémico a pesar de que la creencia resultante sea falsa. Por ejemplo, puede argumentarse que las creencias de Ptolomeo sobre la posición de la tierra con respecto al sol, a pesar de ser falsas, son epistémicamente valiosas, pues él fue un científico meticuloso que hizo todo lo que estuvo a su alcance para alcanzar la verdad sobre la estructura y dinámica del sistema solar. El teórico veritista debería ser capaz de dar una explicación de estos casos.

Claramente, no podemos decir que el valor de este tipo creencias viene de que sean conocimiento, pues son falsas. El siguiente candidato obvio sería la creencia justificada. Las creencias justificadas son consistentes con la falsedad, es decir son un logro epistémico no fáctico. Podríamos decir que Ptolomeo no consiguió conocimiento cuando usaba su modelo planetario, sin embargo sí podemos decir que consiguió creencias justificadas. Si esta fuera la respuesta correcta, entonces el teórico veritista tendría una manera muy directa de explicar el valor epistémico de este tipo de logro en sus términos. Hay que recordar que en la sección (2.1.1) de esta tesis asumí una definición de la justificación epistémica según la cual el valor de ésta es explicable en términos de una teoría veritista. Esta definición sostiene que estar justificado en creer que p es creer que p de una manera tal que es al menos muy probable que la creencia de uno sea verdadera. En términos más precisos, adopté una definición de justificación

⁶⁵ Las razones pueden ser variadas: el modelo puede ser defectuoso, en el sentido de ser no verídico de una manera evidente, el científico que usa el modelo puede ser descuidado, puede haber ocurrido algún error de cálculo, etc.

que implica que formar creencias justificadas es un método fiable de formación de creencias, es decir es un método de formación de creencias que ejemplifica la relación de conductividad a la verdad. Como *ser conducente a la verdad* es una de las relaciones adecuadas con la verdad que presenté en la sección (1.2), el valor de las creencias justificadas es explicable en términos de la teoría veritista que adopto en esta tesis.

Hay dos problemas con una respuesta de este tipo. El primero es que la concepción de la justificación que acabo de esbozar es controvertida y el segundo es que, si adoptamos esta noción de justificación, entonces las creencias de Ptolomeo sobre el sistema solar no estaban justificadas. Su modelo idealizado no lleva de manera fiable a formar creencias, por ejemplo, sobre la ubicación de los cuerpos del sistema solar. De hecho lleva sistemáticamente a creencias falsas, de modo que no podemos caracterizarlas como justificadas. Ambos problemas están conectados: el tipo de característica epistémicamente valiosa que exhiben las creencias de Ptolomeo es precisamente el tipo de característica que algunos filósofos han utilizado para desafiar la concepción de la justificación que adopté antes. A continuación analizaré con detalle qué postura puede adoptar el teórico veritista ante la controversia sobre la noción de justificación epistémica. Espero que la respuesta a esta cuestión proporcione también una respuesta para el caso problemático que ejemplifica el caso de Ptolomeo. Además, en secciones anteriores (especialmente en (2.1.1) y (2.2)) hice uso de la noción de justificación para explicar el valor de algunas formas de entendimiento científico en términos de la teoría veritista. Ahí dije que una discusión detallada de si la noción de justificación epistémica es, a su vez, explicable en términos de la TVEE se llevaría a cabo posteriormente. En este punto ha llegado el momento de llevar a cabo este análisis.

Como ya señalé en la sección (2.1.1) asumir una concepción de la justificación epistémica según la cual formar creencias justificadas es un método fiable de formación de creencias es

controvertido. Hay una variedad de posturas sobre la naturaleza de la justificación epistémica que son incompatibles con esa postura. En particular hay una variedad de teorías según las cuales un sujeto S puede estar justificado en creer que p aun cuando el método mediante el cual S formó la creencia no sea un método fiable. Este tipo de posturas comúnmente se agrupan bajo la categoría de “teorías internistas de la justificación”. Al contrario, la postura sobre la justificación que he asumido en esta tesis es una teoría externista. Ahora bien, una de las tesis centrales de las teorías internistas es que *no es necesario* que la creencia de S en p haya sido formada de un modo fiable para que S esté justificado en creer que p .⁶⁶ El nuevo problema del genio maligno, presentado por primera vez por Cohen y Lehrer (1983) es quizás el argumento maestro a favor de esta tesis internista. Imaginemos que hay dos sujetos (podemos imaginar que uno es la contraparte de otro en un mundo posible similar) que son idénticos en todo lo que respecta a su vida mental: uno de ellos tiene experiencias indistinguibles de las experiencias del otro, parecen recordar las mismas cosas y parecen estar dispuestos a razonar de las mismas maneras. Imaginemos además que ambos sujetos tienen las mismas creencias y que las justifican de una manera similar. Supongamos, además que uno de los sujetos, llamémoslo S_1 , se encuentra en un escenario escéptico, esto es un escenario en el que un genio maligno alimenta al sujeto con experiencias no verídicas. Por otro lado, el otro sujeto, llamémoslo S_2 , no se encuentra en un escenario escéptico, sino en un mundo normal. Esto significa que los procesos de formación de creencias de S_2 son fiables, mientras que los de S_1 no lo son. Esto significa, desde el punto de vista de una teoría externista de la justificación, que las creencias de S_2 están justificadas, mientras que las de S_1 no lo están. Los teóricos internistas sostienen que esta consecuencia es inaceptable. Ellos argumentan que las intuiciones que despiertan en

⁶⁶ Algunos filósofos también ha argumentado en contra de la tesis de que la fiabilidad del método mediante el cual S formó su creencia de que p sea una condición suficiente para que S esté justificado en creer que p . Bonjour ha argumentado de esta forma en su artículo “Externalist Theories of Epistemic Justification” (1980).

nosotros este tipo de casos muestran que las creencias de S_I sí están justificadas, a pesar de no haber sido formadas mediante un método fiable. Esto significa, argumentan los internistas, que es incorrecta la tesis externista de que una condición necesaria para la justificación de las creencias, es que éstas se formen mediante un proceso fiable de formación de creencias.

Si los internistas están en lo correcto en esta crítica, esto significa que la concepción de la justificación epistémica que he adoptado en esta investigación no es correcta. Esto significa que la explicación que aquí he dado del valor de la justificación epistémica en términos de la TVEE descansa en un entendimiento incorrecto de la noción de justificación epistémica. En este punto el teórico veritista podría adoptar alguna de las siguientes dos alternativas: 1) argumentar que la crítica internista es, en algún sentido, insatisfactoria o 2) argumentar que también el valor epistémico de la justificación entendida del modo internista puede explicarse en términos de la teoría veritista. Argumentaré que lo más razonable para el veritista es adoptar (2).

Muchos epistemólogos externistas han argumentado que el nuevo problema del genio maligno no demuestra que la concepción externista de la justificación epistémica sea errónea. Por ejemplo, Goldman ha propuesto al menos dos soluciones a este problema, que son consistentes con el externismo. La primera de ellas consiste en sostener que la noción de justificación epistémica debe aplicarse sólo a “mundos normales”, donde los mundos normales excluyen a los mundos en los que los sujetos son víctimas de un genio maligno. De manera más precisa: “La creencia de S en p está justificada sólo si el proceso que produjo la creencia de S es fiable en mundos *normales*.” (Goldman, 1986, p.107). De acuerdo con Goldman, los mundos normales son aquellos mundos en los que nuestras creencias cotidianas sobre el mundo son verdaderas: creencias sobre los objetos que existen y sobre los eventos y cambios que ocurren (Cfr., Goldman, 1988, p. 61). Así, Goldman concede que las creencias de S_I sí están justificadas, sin embargo rechaza la conclusión del internista de que esto significa que la fiabilidad no es una condición necesaria para la justificación. De acuerdo con Goldman, las

creencias de S_1 están justificadas porque los procesos que utilizó para formar sus creencias *serían* fiables en un mundo normal.

La segunda solución de Goldman al nuevo problema del genio maligno y que es consistente con el externismo se basa en trazar una distinción entre justificación epistémica fuerte y justificación epistémica débil. La justificación epistémica fuerte está definida del modo externista (*Cfr.*, Goldman, 1988, p. 54). Por otro lado, la justificación epistémica débil se define como sigue: un sujeto S está débilmente justificado en creer que p si su creencia es el producto de un proceso no fiable de formación de creencias y S no tiene la culpa de creer que p . Donde “no tener la culpa de creer que p ” significa que el sujeto no tiene evidencia para creer que su creencia es falsa, que su proceso de formación de creencias no es fiable ni está a su alcance algún modo de determinar que el proceso no es fiable. (*Cfr.*, Goldman, 1988, p. 59). Goldman sostiene que lo que sucede en el caso que se presenta en el nuevo problema del genio maligno es que S_1 está justificado débilmente en tener sus creencias, mientras que S_2 está justificado fuertemente en tener sus creencias. Un externista como Goldman pretende que su noción de justificación explique correctamente sólo los casos en los que hay justificación fuerte. De este modo, el caso problemático de S_1 no representa un problema para el análisis externista de la justificación epistémica y no demuestra que la fiabilidad no sea una condición necesaria para la justificación fuerte, que en última instancia, es la que le interesa al externista.

He presentado estas dos alternativas para ilustrar que hay estrategias abiertas que podría adoptar un epistemólogo externista para rechazar la concepción internista de la justificación epistémica. Hay una variedad de alternativas, además de las que acabo de presentar, que se pueden perseguir si se desea defender una concepción externista sobre una concepción internista de la justificación epistémica. Sin embargo, no es claro que ésta sea la postura más viable para quien desea explicar el valor de la justificación epistémica en términos de una teoría veritista. En primer lugar, adoptar esa estrategia lleva al veritista a enfrascarse en un debate

muy complejo sobre la justificación epistémica del que no es seguro que salga victoriosos. Por otro lado, aún cuando se lograra demostrar que la concepción correcta de la noción de justificación epistémica es la externista, seguiría estando presente la intuición de que hay algo epistémicamente valioso en las creencias de S_I aunque no estén epistémicamente justificadas. Un ejemplo que he presentado anteriormente puede poner de relieve esta cuestión. Recordemos que las creencias que Ptolomeo extraía de la aplicación de su modelo planetario eran falsas. Por ejemplo, la creencia de que el Sol orbita alrededor de la Tierra es falsa. Además, el método de formación de creencias de Ptolomeo *no* era un método fiable. Sin embargo, parece intuitiva la idea de que sus creencias tenían algún valor epistémico. Por lo tanto, el veritista debe explicar en qué consiste ese valor epistémico, independientemente de si las creencias están justificadas o no.

Así, parece que lo más razonable para el teórico veritista que busca dar una respuesta a los epistemólogos que favorecen una concepción internista de la justificación consiste en explicar el valor de la justificación así entendida en términos de la teoría veritista. Haciendo uso de la terminología de Goldman diremos que la tarea del veritista en este punto consiste en explicar el valor epistémico de la justificación débil en términos de una teoría veritista. No es obvia la manera en la que podría explicarse su valor apelando a alguna de las tres relaciones con la creencia verdadera que propuse en la sección (1.2). Es claro que las creencias justificadas débilmente no pueden ejemplificar la relación de ser conducentes a la verdad ni la de ser altamente probables dada la evidencia. Esto queda demostrado claramente si se recuerda el caso que presenta el nuevo problema del genio maligno. En ese caso las creencias de S_I no provenían de un método fiable de formación de creencias y es complicado determinar si la evidencia a favor de esas creencias las vuelve altamente probables. De igual modo, no es claro que el valor de las creencias justificadas débilmente dependa de otras creencias verdaderas. En última instancia, la explicación más directa del valor de este tipo de justificación consiste en

que el sujeto ha hecho todo lo que está a su alcance para formar creencias verdaderas. Es claro que su valor está relacionado de alguna manera con la creencia verdadera, sin embargo no parece que esta relación pueda explicarse en términos de las tres relaciones que he venido usando a lo largo de esta tesis. Por lo tanto, en este punto me gustaría sugerir que hace falta añadir una nueva relación para dar cuenta del valor epistémico de este tipo de justificación. Esta relación podría formularse de la siguiente manera: una creencia ejemplifica la relación de ser producto de un método que, desde un punto de vista intersubjetivo, es la mejor manera de alcanzar creencias verdaderas.⁶⁷

Hay que recordar que en varias ocasiones señalé que el veritismo debe ser tratado como una hipótesis que se encuentra constantemente a prueba. No debe resultar sorprendente que en este punto nos hayamos encontrado con un valor epistémico que exige una consideración más seria y, quizás, la postulación de una nueva relación con la creencia verdadera.

Me gustaría finalizar este capítulo retomando el tema principal. Hasta aquí he argumentado que la propuesta de Elgin sobre cuál es la naturaleza del tipo de entendimiento que surge del uso de modelos idealizados es insatisfactoria por varias razones. He propuesto, en su lugar, una postura según la cual este tipo de entendimiento consiste fundamentalmente en un tipo de conocimiento. La diferencia radical entre la propuesta de Elgin y la mía está en el tipo de teoría de la representación que asume cada una. Una de las consecuencias más relevantes de esta diferencia radica en que la teoría de la representación que yo adopto *permite* que las creencias que alcanzamos usando modelos idealizados constituyan conocimiento, mientras que esto es imposible para la propuesta de Elgin. Propuse, además, que es posible explicar el valor de cierto tipo de creencias falsas que surgen de algunos modelos idealizados apelando al valor de la justificación epistémica que parece ser, a su vez, explicable en términos de una teoría

⁶⁷ Fernández (2011, pp.163-164) da una explicación muy parecida a esta cuando explica en qué consiste el valor de una creencia que ha sido formada sin violar ninguna obligación epistémica.

veritista. Sin embargo, una consideración más cuidadosa de esta última cuestión nos llevó a considerar concepciones alternativas de la noción de justificación epistémica que presentaban problemas para el veritista. En particular, nos llevó a considerar algunos tipos de concepciones internistas de la justificación. Señalé que hay al menos dos estrategias generales *viabes* que podría seguir el veritista para dar cuenta de las inquietudes internistas. No argumenté de manera concluyente que ninguna de las dos pudiera dar una respuesta favorable para el veritista. Sin embargo sí esboqué de qué manera podría darse una respuesta de este tipo. En particular propuse que quizás el veritismo se vería obligado a proponer una nueva relación adecuada con la creencia verdadera para explicar el valor epistémico de la justificación entendida del modo internista. Incluso presenté un esbozo de esta relación. La consecución de cualquiera de estas estrategias queda fuera de los alcances de esta tesis por cuestiones de espacio. Sin embargo sí me gustaría señalar que aunque no se ha logrado establecer concluyentemente que es posible superar las dificultades que presenta una concepción internista de la justificación epistémica, sí se ha logrado llevar la discusión general sobre el valor del entendimiento que surge del uso de modelos idealizado a una discusión mucho más familiar y desarrollada en la epistemología, es decir a la discusión sobre la justificación epistémica.

Conclusiones

Hay varias cuestiones importantes que han surgido a lo largo de esta investigación. En primer lugar, he mostrado aquí que es posible explicar en términos de una teoría veritista el valor de una amplia variedad de formas de entendimiento científico. Hay que recordar que en el capítulo 1 insistí en la idea de que la tesis veritista que adopto aquí tiene el estatus de hipótesis. Esto quiere decir que el apoyo a su favor debe provenir de su capacidad de explicar una amplia variedad de logros epistémicos en sus términos. Por lo tanto, esta investigación aporta evidencia a favor de la idea de que la tesis veritista de evaluación epistémica es correcta.

Resumiendo los resultados de estos análisis, se mostró que todas las formas de entendimiento científico analizadas en esta tesis ejemplifican alguna de las relaciones adecuadas con la creencia verdadera que se propusieron en la sección (1.2). El tipo

entendimiento científico que surge de las explicaciones científicas del tipo N-D e I-E ejemplifica claramente las relaciones de dependencia, de conductividad a la verdad y el de ser altamente probable dada cierta evidencia. Tanto los argumentos que constituyen a una explicación N-D e I-E deben tener premisas verdaderas o, cuando menos, epistémicamente justificadas, de modo que el entendimiento científico que se alcanza utilizando este tipo de explicaciones ejemplifica la relación de dependencia. Además, las explicaciones N-D arrojan conclusiones verdaderas y las explicaciones I-E arrojan conclusiones altamente probables, dada la evidencia presente en las premisas. De modo que el entendimiento científico que surge de estas explicaciones ejemplifica las relaciones de conductividad a la verdad y de ser altamente probable, dada la evidencia. El tipo de entendimiento científico que surge de las explicaciones por relevancia estadística y el que surge de las explicaciones causales ejemplifica la relación de dependencia con la creencia verdadera, pues en ambos casos una buena explicación depende crucialmente de que cierta información sea verdadera. El tipo de entendimiento científico que surge de la incorporación teórica de un fenómeno ejemplifica la relación de dependencia con la creencia verdadera y la de ser altamente probable, dada la evidencia. Los argumentos en sentido amplio juegan un papel muy relevante en la incorporación teórica. Un argumento en sentido amplio es una estructura inferencial que incrementa la probabilidad de su conclusión a un grado “suficientemente” alto. Así, el entendimiento que surge de la incorporación ejemplifica la relación de dependencia. Por otro lado, un requisito adicional para que haya una incorporación exitosa, de acuerdo con esta teoría, consiste en que los elementos del corpus cognitivo al que se incorpora el fenómeno no estén conectados circularmente y, además, que haya unificación con los datos empíricos. Argumenté que el valor de un corpus cognitivo de este tipo depende de que las creencias sobre los datos empíricos, que constituyen la base de todo el corpus, estén epistémicamente justificados. Lo cual significa que el entendimiento que surge de la incorporación de un fenómeno a un corpus cognitivo también ejemplifica la relación de

dependencia con la creencia verdadera. Finalmente se mostró que el tipo de entendimiento científico que surge del uso de modelos idealizados ejemplifica la relación de conductividad a la verdad. Argumenté en el capítulo 3 que el entendimiento que surge a partir del uso de modelos idealizados en la actividad científica consiste, fundamentalmente, en conocimiento. De modo que razonar haciendo uso de modelos idealizados ejemplifica la relación de ser conducente a la verdad, pues de aplicarse correctamente el uso de modelos idealizados nos lleva a alcanzar conocimiento, que es un logro epistémico fáctico.

Ahora bien, la justificación epistémica tuvo un papel importante en la explicación del valor de varias formas de entendimiento científico. En particular fue útil para explicar por qué es valiosa una manera de entender el entendimiento que surge del uso de explicaciones N-D e I-E. De acuerdo con esta manera de entender estas explicaciones, no hace falta que las premisas de las explicaciones N-S e I-E sean verdaderas, sino que basta con que estén justificadas. Cuando expliqué el valor del entendimiento que surge de la incorporación teórica de un fenómeno a un corpus cognitivo también hice uso del valor de la justificación de las creencias en datos empíricos. Finalmente y de manera más sobresaliente, el tema de la justificación epistémica volvió a aparecer cuando se intentó explicar cuál es el valor epistémico de algunas creencias falsas que se adquieren haciendo uso de modelos idealizados.

Sin embargo, como señalé al final de la de la sección (3.6), las concepciones internistas de la justificación epistémica son particularmente difíciles de explicar desde un punto de vista veritista. En esa sección presenté dos estrategias que podría adoptar el teórico veritista. La primera consistía en argumentar en contra de las concepciones internistas y a favor de alguna concepción externista de la justificación epistémica. La segunda consistía en argumentar que incluso el valor de la justificación epistémica entendida de un modo internista es explicable en términos de una teoría veritista. Sostuve que, a pesar de que ambas estrategias son viables desde el punto de vista de la epistemología contemporánea, la segunda estrategia es más

atractiva para el veritista, pues adoptar esta estrategia le permite mantenerse neutral en el debate internismo/externismo sobre la justificación epistémica. Además, si el veritista eligiera tomar el lado del externismo en el debate sobre la justificación epistémica, se podría argumentar que aún debe explicar cuál es el valor epistémico de aquello que los internistas erróneamente identifican con la justificación epistémica. Señalé, sin embargo, que no es claro que sea posible explicar el valor de esta concepción de la justificación epistémica apelando sólo a las tres relaciones que presenté en la sección (1.2). Lo cual significa que quizás el veritismo se vería obligado a proponer una nueva relación adecuada con la creencia verdadera para explicar el valor epistémico de la justificación entendida del modo internista. No parece que tal empresa sea imposible, pues después de todo es claro que una creencia justificada internistamente es valiosa porque el sujeto que la forma ha hecho todo lo que está a su alcance para formar una *creencia verdadera*. Por esta razón propuse, a manera de sugerencia, una relación de la que podría echar mano el veritista para explicar su valor: que sea producto de un método que, desde un punto de vista intersubjetivo, es la mejor manera de alcanzar creencias verdaderas. Pero esta es meramente una elucidación esquemática de la manera en que el veritismo podría explicar el valor de la justificación epistémica entendida del modo internista. Una explicación completa y satisfactoria de que esto es así requeriría de un análisis exhaustivo y detallado de distintas teorías internistas de la justificación. En este momento no cuento con el tiempo para llevar a cabo semejante tarea. A pesar de que en este punto dejo la investigación abierta, me gustaría señalar que al menos se ha logrado reducir una serie de cuestiones relativas al valor de distintas formas de entendimiento a una discusión sobre el valor de una noción mucho más familiar en la epistemología contemporánea, me refiero a la concepción internista de la justificación epistémica.

Esto en lo que respecta al debate sobre el veritismo. Hay, sin embargo, una vertiente de esta investigación que también es pertinente para el debate sobre la naturaleza del

entendimiento. Como expliqué en la introducción de esta tesis, hay dos posturas principales sobre la naturaleza del entendimiento: la primera, enarbolada prominentemente por algunos epistemólogos, sostiene que el entendimiento es un logro epistémico distinto del conocimiento; la segunda, enarbolada típicamente por algunos filósofos de la ciencia, sostiene que el entendimiento es simplemente un tipo de conocimiento, en particular conocimiento de causas. Elgin pertenece a la primera clase de filósofos y sostiene, en particular, que el entendimiento no es un tipo de conocimiento, pues según ella uno puede tener entendimiento sin tener conocimiento. La razón más fuerte de Elgin para sostener su postura consiste en señalar que el tipo de entendimiento asociado al uso de modelos idealizados no es fáctico. De modo que este tipo de entendimiento no puede ser un tipo de conocimiento, pues mientras el primero no es fáctico, el segundo sí lo es. Señalé que muchos de los supuestos de Elgin son incorrectos y, además, argumenté que hay varios problemas relacionados con su propuesta particular. Esto significa, en primer lugar, que las razones de Elgin para argumentar que este tipo de entendimiento es no fáctico son incorrectas. Como consecuencia, no parece haber ningún obstáculo para adoptar la postura de que este tipo de entendimiento no es sino una forma de conocimiento, no necesariamente conocimiento de causas, pero sí un tipo de conocimiento.

Así, mi propuesta consiste en señalar que el entendimiento que surge de este tipo de modelos *puede* ser conocimiento. Quizás Elgin es quien ha presentado los argumentos más fuertes en contra de la idea de que el entendimiento sea un tipo de conocimiento y en contra de que el entendimiento sea un logro epistémico fáctico. Al menos esta línea argumentativa ha sido desmantelada aquí. De modo que la aportación de esta tesis al debate sobre la naturaleza del entendimiento consiste en eliminar un obstáculo para las teorías que pretenden identificar al entendimiento con una forma de conocimiento o con un logro epistémico fáctico.

De manera abierta Stephen Grimm, Jonathan Kvanvig y Duncan Pritchard han argumentado a favor de la tesis de que sea lo que sea el entendimiento, es un logro epistémico

fáctico. Sus razones para sostener esta tesis consisten en ejemplos cotidianos en donde se usa el término “entender” de un modo fáctico. Grimm presenta el siguiente ejemplo: imaginemos que un sujeto llega a su casa y se da cuenta de que la luz de su refrigerador no funciona. Mete la mano y se da cuenta de que está a temperatura ambiente. Mientras investiga a qué se debe esto se da cuenta de que el cable está desconectado. En este punto el sujeto podría pensar que *entiende* porqué el refrigerador dejó de funcionar, a saber porque no está conectado. Sin embargo imaginemos que el refrigerador no funciona porque hubo un cortocircuito y que está desconectado para evitar algún accidente. En este caso, a pesar de que el sujeto crea que entiende, afirma Grimm, nuestra intuición nos dice que el sujeto *no* entiende, pues su supuesto entendimiento es falso: el sujeto se representa de un modo no verídico al mundo (*Cfr.*, Grimm 2006, pp. 517-518).⁶⁸ Es cierto que este ejemplo no es un caso de entendimiento científico, sino de entendimiento cotidiano. Sin embargo, Grimm extrae de la consideración de este ejemplo la conclusión general de que el entendimiento de los fenómenos naturales *es fáctico*. No profundizaré en la cuestión de si las razones que presentan Grimm y Kvanvig pueden generalizarse para el caso del entendimiento científico. Aquí sólo quiero ilustrar que en la discusión sobre la naturaleza del entendimiento algunos epistemólogos han aportado razones a favor de la idea de que el entendimiento es un logro epistémico fáctico, lo cual es consistente con la postura que he defendido aquí.

Hay una última cuestión muy relevante en la discusión sobre la naturaleza del entendimiento que me gustaría abordar aquí. De acuerdo con algunos epistemólogos, una característica distintiva del entendimiento es que, a diferencia del conocimiento, aquél implica

68 Kvanvig (2011), pp. 129-130 también argumenta a favor de la facticidad del entendimiento. También Duncan Pritchard (2009) argumenta a favor de una teoría según la cual el entendimiento es un logro epistémico fáctico. Sin embargo, las razones que Pritchard presenta a favor de esta tesis se dirigen sólo en contra de la postura que presenta Zagzebski (2001) y de hecho señala que Elgin ha presentado buenos argumentos a favor de la no facticidad de un tipo de entendimiento científico. Sin embargo, Pritchard no le dedica espacio a un análisis de las razones que presenta Elgin.

que el sujeto que entiende tiene además algunas capacidades intelectuales asociadas con el entendimiento en cuestión. Por ejemplo, de acuerdo con Grimm, entender implica tener la habilidad para responder preguntas del tipo “¿Qué hubiera ocurrido si las cosas hubieran sido diferentes?”⁶⁹ Una de las consecuencias importantes de esta concepción sobre el entendimiento es que no importa cuánto conocimiento de un tema tenga un sujeto, si no tiene esta habilidad, entonces no tiene entendimiento.⁷⁰ La investigación que he llevado a cabo en esta tesis no proporciona elementos para tomar una postura frente a esta cuestión. Mi postura deja abierta esta cuestión porque a lo más implica algo sobre la relación del entendimiento con el conocimiento *proposicional*, pero no implica nada sobre la relación entre aquél y el saber cómo (*knowing how*). Es cierto que el tipo de entendimiento que surge de las explicaciones causales (como las entiende Woodward) bien podría tener asociada una habilidad como ésta. Sin embargo, una respuesta satisfactoria a cuál es la relación entre el entendimiento científico y estas habilidades intelectuales requiere de una investigación profunda y detallada que no se ha llevado a cabo aquí. Por lo tanto, queda abierta la cuestión de cuál es exactamente esta relación.

Recapitulando, podríamos extraer las siguientes conclusiones de esta investigación: 1) el valor de las distintas formas de entendimiento científico que se analizaron aquí es explicable en términos de una teoría veritista, 2) se propuso a manera de esbozo una forma en la que podría explicarse el valor de la justificación epistémica entendida del modo internista en términos de una teoría veritista, 3) las razones de Elgin para sostener que el tipo de entendimiento científico que surge del uso de modelos idealizados es un logro epistémico no fáctico, son bastante cuestionables, 4) el logro epistémico asociado con el uso de modelos idealizados consiste, fundamentalmente, en un tipo de conocimiento y, finalmente, 5) la postura que se adopta aquí

⁶⁹ Grimm (2006), p. 18. Esta idea, según el propio Grimm, proviene de la teoría contrafáctica de la explicación causal que defiende James Woodward (2003) y que presenté aquí en la sección (2.1.3).

⁷⁰ Pritchard parece sostener una tesis parecida en su “Knowledge, Understanding and Epistemic Value” (2009).

sobre cuál es el logro epistémico asociado al uso de modelos idealizados no proporciona elementos para pronunciarse sobre la idea de si tener entendimiento implica tener ciertas habilidades.

Bibliografía

Achinstein, Peter. 1983. *The Nature of Explanation*. Nueva York: Oxford University Press.

Alston, William P. 1989. "A Doxastic Practice Approach to Epistemology" en *Knowledge and Scepticism*. Clay, M. & Lehrer, K. (Eds.) Boulder, Colorado: Westview.

Alston, William P. 2005. *Beyond "Justification"*. Ithaca: Cornell University Press.

Barberousse, Anouk y Ludwig, Pascal. 2009. "Models as Fictions" en *Fictions in Science. Philosophical Essays on Modeling and Idealizations*. Mauricio Suárez (Ed.) Nueva York: Routledge.

BonJour, Laurence. 1980. "Externalist Theories of Empirical Knowledge" en *Midwest Studies in Philosophy* 5 (1): 53-73.

- BonJour, Laurence.** 1985. *The Structure of Empirical Knowledge*. Cambridge, Massachusetts: Harvard University Press.
- Cartwright, Nancy.** 1983a. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, Nancy.** 1983b. “Do the Laws of Physics State the Facts” en su *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, Nancy.** 1983c. “When Inference Leads to Explanation” en su *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Chignell, Andrew.** 2010. “The Ethics of Belief” en *The Stanford Encyclopedia of Philosophy* (edición del otoño de 2010), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2010/entries/ethics-belief/>.
- Cohen, L. Jonathan.** 1992. *An Essay on Belief and Acceptance*. Nueva York: Oxford University Press.
- Cohen, Stewart.** 1984. “Justification and Truth” en *Philosophical Studies* 46 (3): 279-295.
- Cohen, Stewart y Lehrer, Keith.** 1983. “Justification, Truth and Coherence” en *Synthese* 55 (2): 191-207.
- Da Costa, Newton C.A. y French, Steven.** 2003. *Science and Partial Truth. A Unitary Approach to Models and Scientific Reasoning*. Oxford: Oxford University Press.
- Davidson, Donald.** 2001. “A Coherence Theory of Truth and Knowledge” en *Subjective, Intersubjective, Objective*. Nueva York: Oxford University Press.
- Díez, José A. y Moulines, C. Ulises.** 1997. *Fundamentos de Filosofía de la Ciencia*. Barcelona: Editorial Ariel.
- Elgin, Catherine Z.** 1996. *Considered Judgement*. Princeton, Nueva Jersey: Princeton University Press.
- Elgin, Catherine Z.** 2006. “From Knowledge to Understanding” en *Epistemology Futures*. Stephen Hetherington (Ed.). Nueva Yor Oxford University Press.

- Elgin, Catherine Z.** 2007. "Understanding and the Facts" en *Philosophical Studies* 132: 33-42.
- Elgin, Catherine Z.** 2009. "Exemplification, Idealization and Scientific Understanding" en *Fictions in Science. Philosophical Essays on Modeling and Idealizations*. Mauricio Suárez (Ed.) Nueva York: Routledge.
- Fernández Vargas, Miguel Ángel.** 2011. "¿Cómo justificar el veritismo?" en *Dianoia, Vol. LVI, num. 67*. México: UNAM.
- Frigg, Roman y Hartmann, Stephan.** 2012. "Models in Science" en *The Stanford Encyclopedia of Philosophy* (Edición de la primavera de 2012), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2012/entries/models-science/>.
- Goldman, Alvin I.** 1976. "Discrimination and Perceptual Knowledge" en *The Journal of Philosophy* Vol. 73, No. 20, pp. 771-791.
- Goldman, Alvin.** 1986. *Epistemology and Cognition*. Cambridge, Massachusetts: Harvard University Press.
- Goldman, Alvin I.** 1988. "Strong and Weak Justification" en *Philosophical Perspectives, 2, Epistemology*: 51-69.
- Goldman, Alvin I.** 1999. *Knowledge in a Social World*. Oxford: Clarendon Press; Nueva York: Oxford University Press.
- Goldman, Alvin I.** 2001. "The Unity of Epistemic Virtues" en *Virtue Epistemology. Essays on Epistemic Virtue and Responsibility*. Linda Zagzebsky y Arbor Fairweather (eds.). Nueva York: Oxford University Press.
- Grimm, Stephen R.** 2006. "Is Understanding a Species of Knowledge?" en *British Journal for the Philosophy of Science* 57 (3):515-535.
- Grimm, Stephen R.** 2008. "Epistemic Goals and Epistemic Values" en *Philosophy and Phenomenological Research* Vol. LXXVII, No. 3.

- Hempel, Carl Gustav y Oppenheim, Paul.** 1948. "Studies in the Logic of Explanation" en *Philosophy of Science*. Vol. 15, No. 2 (Abril), pp. 135-175.
- Hempel, Carl Gustav.** 1965. "Aspects of Scientific Explanation" en *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Nueva York: Free.
- Hempel, Carl Gustav.** 1966. *Philosophy of Natural Science*. Nueva Jersey: Prentice-Hall.
- Kaplan, David.** 1989. "Demonstratives" en *Themes from Kaplan*. Joseph Almog, John Perry y Howard Wettstein (Eds.). Nueva York: Oxford University Press.
- Kvanvig, Jonathan.** 2011. "Conocimiento y entendimiento" en *Normas, virtudes y valores epistémicos*. Margarita M. Valdés y Miguel Ángel Fernández (Eds.). Traducción por Diego Ortiz González. México: UNAM.
- Lapp, Ralph E.** 1969. *Materia*. Hamburgo: Time-Life.
- Laudan, Larry.** 1984. "Explaining the Success of Science: Beyond Epistemic Realism and Relativism" en *Science and reality: Recent Work in the Philosophy of Science. Essays in Honour of Ernan McMullin*. Gary Gutting, et al. (eds.). Notre Dame, Indiana: University of Notre Dame Press.
- Lipton, Peter.** 2004. *Inference to the Best Explanation, Second edition*. Nueva York: Routledge.
- Lipton, Peter.** 2009. "Understanding without Explanation" en *Scientific Understanding: Philosophical Perspectives*. Henk W. de Regt, Sabina Leonelli y Kai Eigner (Eds.) Pittsburgh: University of Pittsburgh Press.
- McMullin, Ernan.** 1985. "Galilean Idealization" en *Studies in History and Philosophy of Science* Vol. 16, no. 3: 247-273.
- Mikenberg, Irene; Da Costa, Newton C. A. y Chuaqui Rolando.** 1986. "Pragmatic Truth and Approximation to Truth" en *The Journal of Symbolic Logic*, Vol. 51 No. 1: 201-221.

- Mizrahi, Moti.** 2011. "Idealizations and Scientific Understanding" en *Philosophical Studies*. DOI: 10.1007/s11098-011-9716-3. Holanda: Springer.
- Perry, John.** 1998. "Indexicals and demonstratives" en *A Companion to the Philosophy of Language*. Bob Hale y Crispin Wroght (Eds.). Oxford: Blackwell.
- Pritchard, Duncan.** 2009. "Knowledge, Understanding and Epistemic Value" en *Royal Institute of Philosophy Supplement 64*.
- Pritchard, Duncan.** 2011. "El Problema del valor del conocimiento" en *Normas, virtudes y valores epistémicos*. Margarita M. Valdés y Miguel Ángel Fernández (Eds.). Traducción por Diego Acosta. México: UNAM.
- Pryor, James.** 2005. "There is Immediate Justification" en *Contemporary Debates in Epistemology*, (M. Steup y E. Sosas, eds.). Malden, Massachusetts: Blackwell.
- Rife, Patricia.** 2007. *Lise Meitner and the Dawn of the Nuclear Age*. Boston, Massachusetts: Virhäuser Verlag.
- Riggs, Wayne D.** 2003 "Understanding 'Virtue' and the Virtue of Understanding" en *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Michael DePaul y Linda Zagzebski (Eds.). Oxford: Oxford University Press.
- Salmon, W.C.** 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, Nueva Jersey: Princeton University Press.
- Schurz, Gerhard y Lambert, Karel.** 1994. "Outline of a Theory of Scientific Understanding" en *Synthese* 101, pp. 65-120. Dordrecht: Kluwer Academic Publishers.
- Suárez, Mauricio** (Ed.). 2009. *Fictions in Science. Philosophical Essays on Modeling and Idealizations*. Nueva York: Routledge.
- Teller, Paul.** 2008. "Representation in Science" en *The Routledge Companion to the Philosophy of Science*. Stasis Psillos y Martin Curd (Eds.). Oxford: Routledge.

- Teller, Paul.** 2009. "Fictions, Fictionalization and Truth in Science" en *Fictions in Science. Philosophical Essays on Modeling and Idealizations*. Mauricio Suárez (Ed.). Nueva York: Routledge.
- Thagard, Paul; Eliasmith, Chris; Rusnock, Paul y Shelley, Cameron.** 2002. "Epistemic Coherence" en *Common sense, reasoning and rationality. Vancouver Studies in cognitive Science Vol. 11*. Oxford University Press.
- Van Fraassen, Bas C.** 1977. "The Pragmatics of Explanation" en *American Philosophical Quarterly*. Vol. 14, No. 2 (Abril), pp. 143-150. Illinois, EUA: University of Illinois Press.
- Van Fraassen, Bas C.** 1980. *The Scientific Image*. Oxford: Clarendon Press.
- Van Fraassen, Bas C.** 2008. *Scientific Representation: Paradoxes of Perspective*. Oxford: Clarendon Press; Nueva York: Oxford University Press.
- Veihinger, Hans.** 2009. *The Philosophy of "As if". A System of the Theoretic, Practical and Religious Fictions of Mankind*. Traducción por C.K. Ogden. Londres: Routledge and Kegan Paul.
- Velleman, J. David.** 2000. "On the Aim of Belief" en *The Possibility of Practical Reason*. Nueva York: Oxford University Press.
- Weisberg, Michael.** 2007a. "Who is a Modeler" en *British Journal for the Philosophy of Science* 58 (2): 207-233.
- Weisberg, Michael.** 2007b. "Three Kinds of Idealization" en *Journal of Philosophy* 104 (12): 639-659.
- Winsberg, Eric.** 2009. "A Function for Fictions: Expanding the Scope of Science" en *Fictions in Science. Philosophical Essays on Modeling and Idealizations*. Mauricio Suárez (Ed.) Nueva York: Routledge.
- Woodward, James.** 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, James. 2011. “Scientific Explanation” en *The Stanford Encyclopedia of Philosophy* (Edición del invierno de 2011). Edward N. Zalta (ed.), URL: <<http://plato.stanford.edu/archives/win2011/entries/scientific-explanation/>>.

Zagzebski, Linda. 2001. “Recovering Understanding” en *Knowledge, Truth, and Duty: Essay on Epistemic Justification, Responsibility, and Virtue*, M. Steup (editor). Nueva York: Oxford University Press.

Zumdahl, Steven S. y DeCoste, Donald J. 2010. *Introductory Chemistry: A Foundation*. Boston, Massachusetts: H Mifflin.