



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN FILOSOFÍA DE LA CIENCIA

CIENCIAS COGNITIVAS

LOS EXPERIMENTOS MENTALES DESDE UNA PERSPECTIVA ABDUCTIVISTA

T E S I S

QUE PARA OPTAR POR EL GRADO DE MAESTRA EN FILOSOFÍA DE LA CIENCIA

PRESENTA:

JÉSSICA GEORGINA CABUTO GARCÍA

ASESORAS:

DRA. ANA ROSA PÉREZ RANSANZ - IIF, UNAM

DRA. FERNANDA SAMANIEGO BAÑUELOS - FFyL, UNAM

MÉXICO, D. F., JUNIO 2015



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

Quiero agradecer al CONACYT, por haberme otorgado una beca desde agosto del 2012 hasta julio del 2014 para realizar mis estudios de maestría en el Instituto de Investigaciones Filosóficas de la Universidad Nacional Autónoma de México y con ello brindarme la oportunidad de enfocarme exclusivamente en mis estudios.

A la UNAM, por darme la oportunidad de estudiar de manera gratuita en una institución de excelencia académica. Haber tenido la oportunidad de formarme en esta sólida comunidad de profesores e investigadores ha sido la mejor experiencia.

A mis asesoras, la Dra. Ana Rosa Pérez Ransanz y la Dra. Fernanda Samaniego Bañuelos por todo el apoyo que durante esta etapa me brindaron, por dedicarme siempre un espacio para discutir mi trabajo, por sus buenos consejos, comentarios y excelente orientación para sacar adelante este proyecto. Les agradezco profundamente haberme dado la oportunidad de conocerlas académica y personalmente.

A mis sinodales, la Dra. Atocha Aliseda Llera, el Dr. Jorge Roberto Ornelas Bernal y el Dr. Matthieu Fontaine, por haber aceptado leer mi trabajo, por todas sus observaciones, críticas y objeciones que enriquecieron de manera significativa el mismo y por alentarme a buscar mejores formas no sólo de expresarme, sino de ver las cosas, explorando diferentes perspectivas y ampliando cada vez más el panorama.

A todos mis profesores que tuve durante la maestría, por haberme presentado los contenidos de cada curso de manera estimulante y proporcionarme las herramientas necesarias para ejercer de mejor manera mi profesión. Muchas gracias por todos los debates, discusiones y reflexiones dentro y fuera del aula.

A la Coordinación del Posgrado y al Comité Académico por darme la oportunidad de ingresar a la maestría y por todo su apoyo en los procesos administrativos.

Agradezco el apoyo que me brindaron para poder realizar actividades académicas que me enriquecieron profesionalmente, en particular la oportunidad de poder asistir a un congreso internacional y poder realizar una estancia de investigación en el extranjero.

A mi familia, por haberme enseñado que el estudio es la mejor forma de invertir mi tiempo y esfuerzo. Gracias por alentarme siempre a seguir preparándome y por todo su apoyo que a pesar de la distancia me hicieron sentir constantemente.

A mi esposo José Luis, por interesarse en mi trabajo, compartirme sus ideas y reflexiones, por desvelarse a mi lado y por alentarme a seguir superándome como profesional y como persona. Su amor incondicional es mi inspiración y fortaleza en este proyecto de vida que decidimos compartir. Gracias ma, por estar a mi lado apoyándome siempre.

Gracias a todos, los que con su amor, apoyo y buenos consejos hicieron posible concretar este objetivo.

Agradecimientos	i
Índice	iii
Introducción	1
1. Algunas características de los experimentos mentales.	4
1.1 Cualidad <i>a priori</i> .	4
1.2 Elemento modal y carácter contrafáctico.	8
1.3 Función heurística.	14
1.4 Característica metodológica.	17
2. Sobre la utilidad de los experimentos mentales.	20
2.1 Apologetas y detractores.	21
2.2 Una consideración desde la modalidad.	23
3. La estructura argumentativa de los experimentos mentales.	29
4. Razonamiento abductivo.	36
4.1 Diferencia entre deducción, inducción y abducción.	39
4.2 Abducción e inferencia a la mejor explicación.	43
5. Aplicación.	46
5.1 El experimento mental del reloj en la caja entre Einstein y Bohr.	47
5.2 El experimento mental de Galileo y las balas de cañón.	51
5.3 El experimento mental de Bell y las naves espaciales.	55
6. Conclusiones.	59
7. Referencias.	62

Introducción

Los experimentos mentales forman parte de la caja de herramientas que el científico o el filósofo utilizan en la investigación. Al consolidarse como parte constitutiva de la metodología científica y filosófica, cambiaron la forma en que tradicionalmente se investigaba y se argumentaba en la ciencia. Por lo que una de las discusiones más interesantes es si éstos involucran estructuras argumentativas y de ser así de qué tipo.

En este trabajo defenderé que los experimentos mentales constituyen herramientas metodológicas útiles en la ciencia, que sirven para fortalecer o debilitar hipótesis o teorías y que algunos experimentos mentales pueden ser reconstruidos como argumentos de tipo abductivo, esto es, como estructuras lógicas constituidas de ciertas premisas que llevan a conclusiones plausibles o no concluyentes.

Analizaré una de las propuestas más conocidas sobre la estructura argumentativa de los experimentos mentales, a saber, la de John Norton (1991). Para Norton los experimentos mentales se pueden reconstruir, principalmente, de manera deductiva, aunque no descarta alguna otra forma de reconstrucción. Las reconstrucciones argumentativas de tipo deductivo son estructuras que constan de ciertas premisas y a partir de las cuales se deriva con necesidad una única conclusión. Mostraremos que esta caracterización causa problemas en ciertos casos, ya que algunos experimentos mentales pueden dar lugar a dos conclusiones distintas y contrarias, como el experimento mental del reloj en la caja entre Einstein y Bohr, el experimento mental de Galileo y las balas de cañón y el experimento mental de las naves espaciales de J. S. Bell¹.

En cambio, desde una perspectiva abductivista es posible dar cuenta de los tres casos anteriores, ya que al considerar a los experimentos mentales como estructuras abductivas, éstos no están obligados a derivar con necesidad una única conclusión o explicación, sino un conjunto de explicaciones posibles y plausibles. Las reconstrucciones abductivas permiten explicar por qué un mismo tipo de experimento mental puede dar lugar a dos o más conclusiones o explicaciones distintas y el hecho de que ciertos experimentos mentales hayan sido aceptados en un determinado momento en el desarrollo de la ciencia y rechazados en otro periodo. Desde esta perspectiva los experimentos mentales son considerados igualmente falibles, pues en

¹ Estos tres casos se desarrollarán a detalle en las secciones 5.1, 5.2 y 5.3 respectivamente.

ocasiones descansan en consideraciones que a la larga se descartan, así como una manera legítima de argumentar en la ciencia.

En el capítulo 1 se presentan, de manera general, algunas de las principales características que las diferentes propuestas sobre la naturaleza de los experimentos mentales han considerado esenciales. El objetivo es mostrar que si bien no existe una única definición acerca de la naturaleza de los experimentos mentales, es posible distinguir algunos elementos comunes o notas distintivas en la pluralidad de concepciones, como la cualidad *a priori*, el elemento modal y contrafáctico, el papel heurístico y sobre todo la función metodológica.

En el capítulo 2 se introduce una de las discusiones más importantes sobre los experimentos mentales, a saber, si éstos constituyen una herramienta útil en la ciencia. Se presenta de manera general la discusión entre apologetas y detractores y se introduce una consideración desde la modalidad que considero arroja luz sobre esta discusión. Mi posición aquí es que los experimentos mentales, al constituir herramientas metodológicas para fortalecer o debilitar hipótesis o teorías, sí constituyen una herramienta útil al servicio de la investigación científica.

En el capítulo 3 se introduce la discusión sobre si los experimentos mentales involucran estructuras argumentativas y de ser así de qué tipo. Se presentan algunas de las principales propuestas que consideran que los experimentos mentales son argumentos o que éstos pueden ser reconstruidos como argumentos. La discusión entre John D. Norton (1991) y Michael Bishop (1999) resulta particularmente interesante para ilustrar este punto, ya que mientras Norton considera que los experimentos mentales involucran estructuras argumentativas de tipo deductivo principalmente, Bishop considera que no es así. Mi intención aquí no es favorecer ningún enfoque reduccionista, sino presentar únicamente la discusión sobre si los experimentos mentales pueden ser reconstruidos como argumentos y de ser así de qué tipo. El objetivo es mostrar que las reconstrucciones deductivas causan problemas, ya que en ciertos casos los experimentos mentales derivan más de una conclusión, en cambio, una reconstrucción abductiva puede dar cuenta perfectamente de este tipo de casos. Mi compromiso es defender que al menos algunos experimentos mentales pueden ser reconstruidos de manera abductiva, lo cual no excluye otras formas de reconstrucción.

El apartado 4 está dedicado al razonamiento abductivo. El objetivo de este capítulo es presentar en qué consiste el razonamiento abductivo, así como en qué se diferencia del razonamiento deductivo, del razonamiento inductivo y de la inferencia a la mejor

explicación, para así poder realizar las reconstrucciones abductivas posteriormente. Señalo la importancia de la noción de plausibilidad y puntualizo que los criterios epistémicos de evaluación y elección dependen de cada comunidad científica. Para desarrollar este capítulo me baso en las propuestas de Charles Sanders Peirce y de la Dra. Atocha Aliseda Llera.

En el apartado 5 se realiza la reconstrucción abductiva de tres experimentos mentales: el experimento mental del reloj en la caja entre Einstein y Bohr, el experimento mental de Galileo y las balas de cañón y el experimento mental de las naves espaciales de J. S. Bell. Y se muestra que los experimentos mentales son una herramienta metodológica que el científico o el filósofo utilizan para fortalecer o debilitar una hipótesis o teoría en la ciencia. En las tres reconstrucciones puede observarse que hay por lo menos dos explicaciones distintas, que fueron plausibles en su momento o como en el caso de Bell, que aún se sigue discutiendo cuál de las dos explicaciones es la más plausible.

Finalmente, cabe mencionar que la propuesta que defiendo no es una propuesta reduccionista, no se trata de reducir a los experimentos mentales a meros argumentos, sino únicamente señalar que algunos de éstos pueden ser reconstruidos como argumentos de tipo abductivo, esto es, como estructuras argumentativas cuyas conclusiones o explicaciones se afirman en términos de plausibilidad. Así mismo, esta propuesta no se excluye otras formas en que los experimentos mentales puedan ser reconstruidos.

Capítulo 1. Algunas características de los experimentos mentales.

El estudio de los experimentos mentales se aborda, principalmente, atendiendo a dos enfoques: el ontológico y el epistémico. El enfoque ontológico se preocupa por la naturaleza de los experimentos mentales, por determinar esencialmente cuáles son sus condiciones necesarias y suficientes para identificarlos. En cambio, el enfoque epistémico se interesa por las cuestiones de conocimiento y justificación, esto es, sobre si los experimentos mentales producen conocimiento y sobre si podemos justificar hipótesis o teorías a través de ellos. Sin embargo, no existe una única propuesta acerca de la naturaleza ontológica y epistémica de los experimentos mentales, sino una pluralidad de concepciones. De esta manera, las características que las distintas propuestas consideran como esenciales son varias y muy diversas.

El objetivo de este capítulo es presentar de manera general algunas de las características más recurrentes entre los interesados en el tema y mostrar que si bien, no existe una única propuesta acerca de la naturaleza de los experimentos mentales, es posible encontrar algunos elementos comunes en esta pluralidad de concepciones, como son la cualidad *a priori*, el elemento modal, el carácter contrafáctico, el papel heurístico y sobre todo la cualidad de ser herramientas metodológicas.

1.1 Cualidad *a priori*.

La cualidad *a priori* ha sido uno de los aspectos más considerados a la hora de intentar explicar qué son los experimentos mentales y cuál es su función en la ciencia. Propuestas como la de James Robert Brown (1991) contemplan su importancia como característica esencial y al mismo tiempo señalan la complejidad del término en sí, ya que éste puede interpretarse de diferentes maneras.

La interpretación semántica del término *a priori* por ejemplo, hace referencia al significado de las palabras, donde el lenguaje desempeña un papel fundamental, ya que la verdad de una proposición dependerá únicamente de los términos involucrados. De esta manera, la frase “*los hombres solteros no están casados*” es verdadera independientemente de la experiencia, su verdad se deduce del significado de las palabras involucradas y no es necesario recurrir a la experiencia para saber que es correcta, ya que el significado de “*soltero*” es “*no casado*”. La experiencia sensorial no

juega un papel importante, ya que la verdad es acerca del lenguaje y no sobre el mundo.

En cambio, desde una interpretación kantiana, se puede considerar que el término *a priori* está más en relación con las condiciones de posibilidad del conocimiento y con la analiticidad e independencia de la experiencia. Para Kant, el conocimiento científico consta, fundamentalmente, de proposiciones o juicios universales y necesarios. Un juicio consiste en la conexión entre dos conceptos, uno de los cuales sirve de sujeto (A) y el otro de predicado (B). Así, identifica tres tipos de juicios:

1. Juicios analíticos: los cuales son proposiciones cuya verdad se desprende directamente de la verdad de sus constituyentes. Aquí el concepto que actúa como predicado (B) está contenido en el concepto que actúa como sujeto (A).
Ejemplo: *Todo cuerpo es extenso*. El concepto “extensión” está contenido en el concepto de “corporeidad”.

El juicio analítico es un juicio que formulamos *a priori*, esto es, sin necesidad de apelar a la experiencia. Con él expresamos de un modo distinto el mismo concepto que expresamos mediante el sujeto; es universal y necesario pero no amplía el conocimiento. La ciencia hace uso de estos juicios para aclarar y explicar un gran número de cosas, pero no se basa en ellos cuando intenta ampliar el conocimiento.

2. Juicios sintéticos: son proposiciones cuya verdad depende tanto del contenido de sus constituyentes como de su relación con el mundo. Aquí el concepto que actúa como predicado (B) puede no hallarse implícito en el concepto que actúa como sujeto (A). El predicado (B) añade al sujeto (A) algo que no se puede establecer por un mero análisis semántico.
Ejemplo: *Todo cuerpo es pesado*. El concepto “pesado” no se obtiene por un mero análisis del concepto “cuerpo”.

El juicio sintético es un juicio que formulamos *a posteriori*, es decir, sólo después de apelar a la experiencia. Este tipo de juicios amplían el conocimiento, ya que siempre nos dicen algo nuevo del sujeto (A), algo que no estaba contenido implícitamente en él. Sin embargo, como dependen de la experiencia no pueden ser universales y necesarios, sino contingentes; a partir de estos juicios solo pueden obtenerse algunas generalizaciones.

3. Juicios sintéticos *a priori*: estos juicios al ser sintéticos amplían el conocimiento del mundo de nuestra experiencia, pero no se basan en la experiencia para su justificación, como los juicios sintéticos *a posteriori*; y al ser *a priori* son

universales y necesarios pero no se basan en el principio de identidad y en el principio de no contradicción como los juicios analíticos *a priori*, ya que no involucran predicados (B) contenidos en el sujeto (A), sino predicados diferentes.

Ejemplo: Todas las operaciones aritméticas, los juicios de la geometría y de la física son sintéticos *a priori*.

Para Kant este tipo de juicios son los juicios propios de la ciencia, ya que unifican la universalidad y necesidad de los juicios analíticos con la posibilidad de ampliar el conocimiento de los juicios sintéticos.

Sin embargo, para Brown la cualidad *a priori* de los experimentos mentales no tiene nada que ver ni con la interpretación semántica ni con la interpretación kantiana, sino más bien con una idea platónica sobre la existencia de los universales -que son propiedades y relaciones-, los cuales tienen una existencia propia muy similar a la de los objetos matemáticos.

According to Platonism, we can intuit some mathematical objects, and mathematical objects are abstract entities. Thus, we can (at least in principle) intuit abstract entities. According to the realist account of laws of nature, laws are also abstract entities. Thus, we might be able (at least in principle) to intuit laws of nature as well. There is one situation in which we seem to have a special access to the facts of nature, namely in thought experiments. Thus, it is possible that thought experiments (at least in some cases) allow us to intuit laws of nature. Intuitions, remember, are nonsensory perceptions of abstract entities. Because they do not involve the senses, they transcend experience and give us a priori knowledge of the laws of nature. Let me head off one concern right away. A priori knowledge is not certain knowledge. Intuitions are open to mistakes, just as ordinary sense perceptions are.

(Brown, 2004; p. 34)

Para Brown la cualidad *a priori* de los experimentos mentales es de suma importancia y hace referencia directa al proceso mental que se lleva a cabo en el pensamiento. A través de los experimentos mentales podemos percibir o tener intuiciones *a priori* de las entidades abstractas, como las leyes de la naturaleza, éstos parten de información empírica pero su proceso es eminentemente *a priori*.

Brown presenta la cualidad *a priori* de los experimentos mentales como un elemento constitutivo, pero también señala que el razonamiento *a priori* es falible y que todo el

conocimiento *a priori* es modificable. De esta manera Brown no se compromete con la infalibilidad del conocimiento *a priori* (Cfr. Brown, 1991; p. 91). También resulta interesante, la taxonomía que propone y que divide, inicialmente, a los experimentos mentales en dos clases: *destructivos* y *constructivos*. Los experimentos mentales constructivos se dividen a su vez en tres tipos: *conjeturales*, *mediativos* y *directos*. Y afirma que existe una clase de experimentos mentales que son simultáneamente *destructivos* y *constructivos*, a éstos los denomina *platónicos* y considera que el papel que juegan en la ciencia es muy especial.

A platonic thought experiment is a single thought experiment which destroys an old or existing theory and simultaneously generates a new one; it is a priori in that it is not based on new empirical evidence nor is it merely logically derived from old data; and it is an advance in that the resulting theory is better than the predecessor theory. (Brown, 1991; p. 76)

Otro autor que también considera el elemento *a priori* en este último sentido es Roy Sorensen (Cfr. 1992; p. 205). Para Sorensen los experimentos mentales también son procedimientos *a priori* que buscan alcanzar su objetivo sin la necesidad de una ejecución física.

Como puede observarse, a pesar de las diferentes interpretaciones que el elemento *a priori* puede tener, se puede decir que en los experimentos mentales esta cualidad ilustra la idea de que éstos son procesos que se llevan a cabo en el pensamiento y que en principio no necesitan ser ejecutados físicamente para alcanzar el mismo resultado. Esto último nos lleva a una de las discusiones más controversiales sobre la naturaleza de los experimentos mentales, a saber, si éstos deben ser en principio estrictamente irrealizables para ser considerados experimentos mentales o si esto no constituye una condición necesaria.

Algunos como M. Cohen (Cfr. 2010; p. 22) consideran que de ser estrictamente irrealizables, esto es, estrictamente *a priori*, los experimentos mentales constituyen una metodología débil que puede tornarse en razonamientos estériles y especulaciones metafísicas. Sin embargo, un gran número de filósofos, matemáticos y físicos, entre otros, consideran que la cualidad *a priori* resulta útil para investigar ciertos aspectos de la realidad que no pueden ser investigados por medios estrictamente físicos y que cuando un experimento mental está inmerso en una teoría y se basa en los datos que ésta proporciona, éste no es una mera especulación ficticia, sino una herramienta metodológica para la evaluación y el fortalecimiento de

hipótesis. Por tanto, los resultados que a partir de éste se deriven serán significativos para la comunidad científica pertinente.

Cualquiera que sea el caso, la cualidad *a priori* es una de las características más destacadas a la hora de intentar dar cuenta de la naturaleza de los experimentos mentales y para los fines de este trabajo se entenderá como la posibilidad de poder ejecutar a los experimentos mentales en el pensamiento, sin descartar la posibilidad de su realización física si así se quiere y las circunstancias físicas lo permiten. La cualidad *a priori* no se entenderá pues, como independencia de la experiencia, sino únicamente con el proceso mental que en principio se puede realizar.

1.2 Elemento modal y carácter contrafáctico.

Otros dos elementos que se consideran importantes a la hora de intentar explicar la naturaleza de los experimentos mentales, son el elemento modal y el carácter contrafáctico. Su presencia y consideración se puede observar desde las primeras definiciones registradas sobre los experimentos mentales. Para Ernst Mach (1948) por ejemplo, los experimentos mentales consisten en la capacidad de imaginar mentalmente una variación de hechos, lo cual implica una exigencia modal, ya que idear un experimento mental involucra crear y concebir un escenario posible con ciertas condiciones específicas para examinar o evaluar algún aspecto de la realidad.

Para Mach esta actividad no es exclusiva del científico, sino también del filósofo, del novelista y del ingeniero.

Aquellos que hacen proyectos, aquellos que construyen castillos [...] romanceros y poetas que se dejan llevar por utopías sociales o técnicas, hacen experimentación mental; la hacen también el comerciante serio, el inventor reflexivo y el sabio. Todos se representan circunstancias diversas y relacionan a estas representaciones ciertas conjeturas. Pero los primeros cambian en su imaginación circunstancias que no se encuentran en la realidad, o bien se representan estas circunstancias seguidas de consecuencias que no tienen vínculos con ellas, mientras que el comerciante, el inventor y el sabio tienen como representaciones buenas imágenes de los hechos y permanecen en sus pensamientos muy próximos a la realidad.

(Mach, 1948; p. 159)

Para Mach los experimentos mentales están en una relación continua con los experimentos del mundo real, éstos últimos son proveedores de datos empíricos y como resultado de esta relación, los experimentos mentales tienen un status epistémico comparable al de los experimentos físicos.

En propuestas más recientes, como la de Pascal Engel (2011) también puede observarse la presencia del elemento modal. Y aunque Engel acepta que no es claro que los experimentos mentales tengan una forma común, aun así sugiere la siguiente estructura:

But it is at least plausible to suggest that thought experiments have the following structure: to devise a thought experiment involves the conception of a possible situation against which we test our intuitions, and from which we reason about an actual case which we contrast with the possible one. So there is, at the basis of a thought experiment, a modal claim to the effect that such and such a situation is possible.

(Engel, 2011; p. 146)

A diferencia del elemento modal, que se refiere al modo en como son las cosas, el carácter contrafáctico se refiere a todo evento o situación hipotética que no ha acontecido en el universo actualmente observable pero que puede ocurrir en un mundo posible². Literalmente es una situación posible “*contraria a los hechos*” que sucede en el universo actual. De esta manera, cuando razonamos a través de un experimento mental estamos contrastando una situación actual con una situación posible. Al igual que el elemento *a priori*, el carácter contrafáctico apela a la noción de posibilidad de ejecución y evaluación en el pensamiento, pero a diferencia de aquel, éste introduce junto con el elemento modal, las posibilidades y la idea de mundos posibles.

Algunas propuestas que han considerado la importancia del carácter contrafáctico son la de Sophie Roux (2011), James. W. McAllister (2012), Timothy Williamson (2007), Pascal Engel (2011) e implícitamente la propuesta de Nancy Nersessian (1992).

² La propuesta de David Lewis sobre los mundos posibles es un realismo modal, desde el cual sostiene que el mundo del que somos parte es uno de una pluralidad de mundos y que nosotros -los que habitamos este mundo- somos sólo unos pocos de todos los habitantes de todos los mundos. Para Lewis el realismo modal se refiere a la pluralidad de mundos, a las diferentes formas en que el mundo puede ser, el mundo actual es sólo una de esas formas. De esta manera, hay tantos mundos como formas distintas en las que un mundo puede ser. Para profundizar más sobre esta propuesta consultar *On the Plurality of Worlds* (1986) de David Lewis.

Para Roux el que los experimentos mentales tengan un carácter contrafáctico, significa que éstos involucran escenarios posibles, los cuales no tienen que ser llevados a cabo de facto para entender o aceptar sus conclusiones. La importancia del carácter contrafáctico se puede observar en las tres características de los experimentos mentales que Roux propone:

1. Son contrafácticos.
2. Involucran un escenario concreto.
3. Tienen una intención cognitiva bien definida.

Cabe mencionar que estas tres características generan tensiones, ya que implican en sí puntos de debate o discusión. Sin embargo, Roux considera que son intrínsecas a la naturaleza de los experimentos mentales y, por tanto, que a través de ellas se ofrece una mejor caracterización.

También resulta interesante la clasificación que Roux desarrolla en torno al concepto de *contrafactividad*. La divide en tres tipos, clases o grados:

1. Contrafactividad débil: indica que la ejecución física de un experimento mental en principio es posible; un experimento mental es débil contrafácticamente cuando es físicamente posible y puede ser llevado a cabo de facto. Independientemente de que se realice o no, la posibilidad de realizarlo está abierta, es decir, que aunque el experimento mental no se haya realizado aún, es posible realizarlo cuando se decida hacerlo.
2. Contrafactividad media: indica la complejidad de realización para un experimento mental, dadas las circunstancias actuales en las que se presenta; un experimento mental tiene contrafactividad media cuando el escenario que éste plantea no solamente no ocurre, sino que pudiese no ocurrir físicamente dadas las condiciones actuales.
3. Contrafactividad fuerte: indica que la ejecución física de un experimento mental no es posible; un experimento mental es fuerte contrafácticamente cuando es físicamente imposible de ser llevado a cabo de facto. La irrealizabilidad se establece desde la teoría en cuestión, es decir, la teoría es la que establece si un experimento es imposible de llevarse a cabo de facto por cuestiones teóricas.

Esta clasificación ayuda a identificar el grado de contrafactividad que implica un determinado experimento mental, ya que si bien todos tienen un carácter contrafáctico, es claro que no todos lo tienen en el mismo grado. Así, habrá experimentos mentales

con un grado de contrafactividad más fuerte o débil que otros y eso determinará en qué medida son realizables.

Como puede observarse, la clasificación también resulta útil para la discusión sobre la realizabilidad³ de los experimentos mentales que ya habíamos comentado. Desde esta perspectiva es posible entender la realizabilidad de manera más amplia, los experimentos mentales no tienen que ser estrictamente irrealizables, sino que algunos de ellos, aquellos que tienen un grado de contrafactividad débil o medio, pueden en principio ser realizables físicamente.

Otro autor que desarrolla e introduce el carácter contrafáctico al intentar dar cuenta de la naturaleza de los experimentos mentales es Williamson (2007). En su propuesta define a los experimentos mentales como una forma de razonamiento contrafáctico, cuya función es mostrar que podría darse un caso determinado donde un estado de cosas actual fuera diferente o simplemente no ocurriese. Para Williamson el razonamiento contrafáctico es una forma de razonamiento ordinario que involucra recursos cognitivos normales como las generalizaciones empíricas y las inferencias. Los experimentos mentales son así, procesos inferenciales en los que utilizamos las mismas habilidades que cuando obtenemos conocimiento del mundo empírico, no interviene nada más que nuestro conocimiento modal de hechos posibles y nuestra capacidad de manejar y evaluar contrafácticos. Esta idea es conocida como la tesis de la contrafactividad de Williamson y cabe mencionar que ésta no afirma que podemos iluminar la naturaleza de los experimentos mentales por trasladarlos a términos contrafácticos, sino únicamente que cuando razonamos a través de experimentos mentales utilizamos los mismos recursos cognitivos que cuando razonamos con contrafácticos.

Algunos como Pascal Engel (2011) consideran que la propuesta de Williamson se basa en la equivalencia entre modalidades y contrafácticos que Stalnaker establece:

(a) „A is necessary“ = „If A were not the case, A would be the case“

$$\Box A = \neg A \Box \rightarrow A$$

(b) 'A is possible' = „It is not the case that if A were the case, A would not be the case“

$$\Diamond A = \neg(A \Box \rightarrow \neg A)$$

³ La realizabilidad de los experimentos mentales se refiere por lo general a que son teóricamente posibles o tecnológicamente posibles.

Y que a partir de estas equivalencias Williamson considera que cuando imaginamos casos posibles a través de experimentos mentales lo único que hacemos es simplemente razonar a través de contrafácticos. Y que así como usamos contrafácticos para razonar sobre posibles estados de cosas, así también podemos usar experimentos mentales para razonar sobre posibles estados de cosas.

Para Engel el carácter contrafáctico de los experimentos mentales se puede observar en que éstos tratan con posibilidades conceptuales, ya que muchos experimentos mentales no se refieren al mundo, sino a nuestros conceptos acerca de él. A diferencia de Williamson que considera que los experimentos mentales tratan con posibilidades metafísicas, Engel considera que a través de los experimentos mentales exploramos los límites y extensión de nuestros conceptos, ya sean ordinarios o filosóficos; no es que los experimentos mentales no se refieran al mundo o a posibles estados de cosas del mundo, sino más bien que nuestra vía de acceso a él es a través de nuestros conceptos.

Algunos otros como McAllister (2013) relacionan el carácter contrafáctico de los experimentos mentales con la imaginación, la cual es entendida como la capacidad mental para concebir entidades, estados de cosas, eventos y fenómenos que no han sido previamente observados. A través de la imaginación se crean contrafácticamente nuevas representaciones mentales.

Y en algunas otras propuestas, como la de Nancy Nersessian (1992) el carácter contrafáctico se señala de manera implícita. Esto se debe a que no todos los interesados en el tema de los experimentos mentales se enfocan en las mismas cuestiones; algunos se concentran en destacar ciertos elementos que consideran poco estudiados e incluso más relevantes, dando por sentado otros elementos y cuestiones. Para Nersessian los experimentos mentales se fundamentan en los modelos mentales y es por ello que se concentra en desarrollar esta idea. Sin embargo, considera de manera implícita el carácter contrafáctico como característica esencial de los experimentos mentales. Veamos.

Para Nersessian los experimentos mentales son una forma de razonamiento basado en la simulación mental. Éstos se presentan como procesos narrativos, ya que cuando razonamos a través de ellos manipulamos mentalmente una situación descrita por medio de una narrativa experimental. Para Nersessian el razonamiento por medio de experimentos mentales implica la construcción de un modelo dinámico en la mente del científico que imagina una secuencia de eventos e infiere resultados.

We can only speculate about what goes on in the mind of the scientist who devises the original thought experiment. Scientists have rarely been asked to discuss the details of how they formulated these experiments. However, reports of thought experiments are presented, customarily, in the form of narratives. Because the thought-experimental narratives are what we have access to and because they are a central form of effecting conceptual change within a scientific community, I propose to examine how the narratives function.

(Nersessian, 1992; p. 296)

La experimentación mental constituye un proceso creativo en la práctica científica, por medio de ésta descubrimos alternativas, hacemos predicciones y obtenemos conclusiones sobre situaciones del mundo real en las que no estamos participando actualmente. El carácter contrafáctico se puede observar en las características que propone de los experimentos mentales, sobre todo en la número 2:

1. Son narrativas: en el momento en que el experimento mental se hace público, éste es presentado en forma de una narrativa y ocasionalmente incluyen ilustraciones visuales. La narrativa llama al público o al lector a imaginar una escena dinámica que se desarrolla en el tiempo y sigue una secuencia causal específica.
2. Invitan al lector a seguir una serie de procesos mentales: el lector es invitado a seguir una secuencia de eventos o procesos como uno lo haría en el mundo real, incluso si la situación puede parecer extraña o fantástica, tal como estar en una caja en el espacio exterior, sin embargo no hay nada de extraño en lo que pasa después, los objetos se comportan como lo harían en el mundo real en la presencia o ausencia de gravedad. La situación es que si el experimento pudiese ser llevado a cabo, la cadena de eventos se desarrollaría de acuerdo con la forma que las cosas suceden en el mundo.
3. Se presentan en su forma pulida o refinada: cuando los experimentos mentales son comunicados públicamente éstos se encuentran en su forma pulida, similar a cuando un experimento físico es publicado. Es decir, no se nos muestra el pensamiento que se llevó a cabo al inventarlo y refinarlo, el lector rara vez escucha los experimentos mentales fallidos o las avenidas exploradas en la construcción de experimento mental.
4. Representan abstracciones: ciertas características de los objetos que podrían estar presentes en el experimento del mundo real no son incluidas, tales como el color de las rocas o las características físicas de los investigadores. Es decir,

ha habido una selección previa de las dimensiones pertinentes sobre las cuales centrar la atención y que evidentemente derivan de nuestra experiencia en el mundo.

5. Es convincente: un experimento mental es, usualmente, tan convincente que incluso en aquellos casos donde es posible realizarlos, el lector no siente la necesidad de llevarlos a cabo.

Si bien, Nersessian no se enfoca específicamente en el carácter contrafáctico, es claro que su propuesta contempla esta idea, ya que manipular mentalmente un escenario hipotético para sacar conclusiones sobre una situación posible, en la cual no estamos participando actualmente, hace referencia directa al carácter contrafáctico, así como al elemento modal.

El elemento modal y el carácter contrafáctico son características fundamentales para una concepción apropiada de los experimentos mentales. Todo experimento mental trata con algún tipo de estructura modal y en ciertos casos con condicionales contrafácticos. En este trabajo el elemento modal se entenderá como referencia directa a la idea de los diferentes tipos de posibilidad (lógica, conceptual, metafísica, etc.), mientras que el carácter contrafáctico se entenderá como referencia a mundos posibles. Al igual que con el elemento *a priori*, ambos elementos apelan a la ejecución mental y evaluación en el pensamiento, sin embargo, aquí tampoco se descarta la posibilidad de su realización física. De hecho, uno de los elementos que más se rescatan sobre este apartado es la clasificación de la contrafactividad que Roux hace, a partir de la cual los experimentos mentales pueden ser catalogados con un grado determinado de contrafactividad y en ese sentido con un grado determinado de realizabilidad.

1.3 El papel heurístico.

Otra de las características más recurrentes en las propuestas sobre la naturaleza de los experimentos mentales es el elemento heurístico. Varios autores han dedicado parte de sus investigaciones a desarrollar y puntualizar esta cuestión pero al igual que el elemento *a priori* el término “heurística” tiene varias interpretaciones en la filosofía de la ciencia.

De manera general las heurísticas se entienden como razonamientos no formales o procesos no rigurosos, que nos lleva a una respuesta o solución, si bien no exacta, muy aproximada a la correcta. Pero para Irme Lakatos (1989) por ejemplo, la heurística es una herramienta para la resolución de problemas que sí involucra técnicas formales, donde la ayuda de técnicas matemáticas sofisticadas permite asimilar las anomalías e incluso las convertirlas en evidencia positiva. Para Lakatos las heurísticas forman parte esencial de los programas de investigación y las clasifica en heurísticas positivas y negativas. Una heurística positiva marca una ruta de investigación auxiliar que debe seguirse, en cambio, una heurística negativa marca las rutas de investigación que deben evitarse, ya que de lo contrario éstas afectarían el núcleo firme de la teoría, es decir, aquel que es “irrefutable” por decisión metodológica de sus defensores.

Otra interpretación del término heurística es la que presenta Gigerenzer (2004), para él la idea de que el buen razonamiento debe cumplir con las leyes de la lógica, el cálculo de probabilidad o la maximización de la utilidad esperada, no es tan cierta como se pretende. Considera que la lógica y las matemáticas son sistemas elegantes pero que no describen la forma real en que las personas actuales razonan y toman decisiones. Por ello, se introduce al estudio de las heurísticas cognitivas, es decir, al estudio de cómo las personas toman decisiones y hacen razonamientos cada día, pues afirma que una buena forma de estudiar las heurísticas es estudiar las experiencias de los sujetos que las implementan. Los resultados de sus estudios muestran que las personas razonan, por lo general, sin usar tanto la lógica o las matemáticas, esto es, sin calcular probabilidades o utilidades.

De esta manera, de acuerdo a Gigerenzer, cuando usamos heurísticas lo hacemos para descubrir buenas soluciones, ya que las soluciones óptimas (aquellas que se encuentran a través de la lógica y matemáticas) están fuera de nuestro alcance. Y así, una heurística puede traducirse en una regla de acción. Sin embargo, una regla de acción no es necesariamente una heurística, a menos que tenga estas tres características:

1. Aprovecha las capacidades del sujeto: una heurística es simple y aprovecha las capacidades del sujeto.
2. Aprovecha la estructura del entorno: la racionalidad de las heurísticas no es lógica, sino ecológica; esto es, la heurística no es buena o mala en sí misma, ni racional o irracional, sino que simplemente aprovecha la estructura del entorno,

es relativa a cada situación y depende de la estructura del ambiente en el cual tiene éxito.

3. Es distinta a la forma de optimización “como-si” característica de los modelos: La optimización se propone para explicar el comportamiento humano, por ejemplo, cuando alguien arroja una pelota y la atrapa, se da por supuesto que se está comportando “como-si” hubiera resuelto un conjunto de ecuaciones diferenciales que predijeran la trayectoria de la pelota. La idea de calcular la trayectoria de la pelota con ecuaciones diferenciales es una forma de optimización. En las heurísticas esto no es así. (Cfr. Gigerenzer, 2004; pp. 63-63).

A partir de estas dos interpretaciones es posible rescatar la idea sobre la heurística que la mayoría de propuestas sobre los experimentos mentales utiliza. El papel heurístico de los experimentos mentales se entiende como una guía para la exploración, indagación o resolución de problemas, que no se basa, en principio, en métodos o técnicas rigurosas y que es falible. Algunos consideran que el hecho de que no se base en métodos o técnicas rigurosas aumenta el grado de su falibilidad. Sin embargo, la falibilidad no es exclusiva de la experimentación mental, la experimentación física tienen sus propias complicaciones y también es falible, solo que en los experimentos mentales el carácter falible es más aceptado y señalado que el que se acepta y señala en los experimentos físicos. No obstante, ambos procesos o tipos de experimentación son falibles, aunque puede considerarse que en grados y formas diferentes, pero al final de cuentas, son ambos falibles.

De esta manera, el papel heurístico en los experimentos mentales hace referencia a que éstos son recursos que nos sirven de guía en la investigación e incluso en la formulación de nuevos principios e intuiciones.

Algunos autores como Thomas Kuhn (1982; p. 37) han dedicado especial atención a la función heurística de los experimentos mentales y han considerado que ésta no sólo sirve como guía en la investigación, sino incluso que su efecto va más allá de eso, ayudando a clarificar conceptos en nuestros esquemas conceptuales que anteriormente eran confusos y obligándonos a reconocer las contradicciones que en un momento dado nuestras teorías implican y las tensiones que nuestros marcos teóricos encierran. Para Kuhn, los experimentos mentales juegan un papel esencial durante los periodos de crisis, ya que al promover el cambio conceptual pueden inducir a los científicos a emplear, como parte integral de su conocimiento, lo que antes les era inaccesible. De esta manera, el papel heurístico de los experimentos mentales

influye no sólo en el desarrollo de la investigación, sino en el desarrollo y construcción de la ciencia misma.

[...] como no implica nueva información sobre el mundo, un experimento mental no puede enseñar nada que no fuera conocido antes. O más bien, no puede enseñar nada acerca del mundo. En cambio, enseña al científico cosas respecto a su aparato mental. Su función se limita a la corrección de anteriores errores conceptuales.

(Kuhn, 1982; p. 37)

Como puede observarse, el papel heurístico de los experimentos mentales constituye otra de sus características máspreciadas, ya que sirve como guía para la formulación de nuevas hipótesis y para la construcción de nuevas formas de explicación. Ésta tiene una función prospectiva y funciona como guía, no pretende ofrecer un algoritmo del descubrimiento, es decir, no pretende ofrecer un conjunto de reglas metodológicas cuya aplicación nos guíe con seguridad al éxito, sino asistirnos o guiarnos, ya sea en la clarificación de conceptos, en la formulación de nuevas hipótesis o en la resolución de problemas, entre otros. Ésta es la interpretación más común en las diferentes propuesta sobre los experimentos mentales y por, tanto, la adoptaremos en esta investigación.

1.4 Característica metodológica.

Los experimentos mentales constituyen una pieza fundamental en la investigación, ya que a través de ellos es posible generar explicaciones plausibles, evaluar, fortalecer o debilitar hipótesis o teorías. Son parte importante de la caja de herramientas de la metodología científica y filosófica y no sólo constituyen una manera legítima de argumentar, sino que incluso cambiaron la forma en que tradicionalmente se investigaba y argumentaba en la ciencia.

Por ello, su función metodológica es fundamental a la hora de intentar dar cuenta de su naturaleza, ya que los experimentos mentales son antes que nada herramientas metodológicas para la investigación, que a su vez pueden ser caracterizados con una función heurística, con un carácter contrafáctico, etc.

Algunos de los autores que se han preocupado por señalar, aunque de manera indirecta, la importancia de la característica metodológica, son James Robert Brown (1986; pp. 1-16) y Alexander Koyré (1968; pp. 75, 88). Para estos autores, el siglo XVII marcó una diferencia en la forma en que tradicionalmente se investigaba en la ciencia. Ambos afirman que la observación empírica, que es uno de los aspectos que caracterizaba la manera tradicional de investigación científica no se incrementó, sino que disminuyó considerablemente. De acuerdo con Brown y Koyré, los experimentos mentales comenzaron a ganar terreno en el campo de la investigación y los científicos comenzaron a utilizarlos cada vez más en sus investigaciones, modificando con ello la forma en que tradicionalmente se investigaba o al menos la forma aceptada en que tradicionalmente se investigaba en la ciencia.

Para Brown, quizás más que cualquier otra cosa, los experimentos mentales representaron una forma diferente de proceder e investigar en la ciencia, mostrando ser herramientas metodológicas de gran valor para la investigación. Mientras que para Koyré éstos demostraron un aspecto importante de la naturaleza del conocimiento, su dimensión sintética *a priori*. Los experimentos mentales son, actualmente, parte esencial en la metodología científica y filosófica; distintas áreas del conocimiento como la física, las matemáticas, la filosofía de la mente, la epistemología, la ética, entre otras, emplean experimentos mentales en sus investigaciones, ya sea para explorar, fortalecer o debilitar determinadas hipótesis, así como para argumentar a favor o en contra de una determinada teoría. Y ha resultado que en ciertas áreas, como la física cuántica y la relatividad, el papel que los experimentos mentales han desempeñado ha sido crucial no sólo para el planteamiento de las teorías, sino para el desarrollo de las teorías mismas.

Sin embargo, al considerar que los experimentos mentales constituyen herramientas metodológicas se derivan diferentes cuestiones respecto a su fiabilidad y su utilidad, entre otras. Es decir, si los experimentos mentales constituyen herramientas metodológicas para la investigación, cabe preguntarse si son herramientas fiables, qué tan fiables son o qué condiciones deben satisfacer para aumentar su fiabilidad. Y de acuerdo a las respuestas que se den a las preguntas anteriores, se responderá en gran medida las cuestiones sobre su utilidad, porque si de entrada se considera que los experimentos mentales no son nada fiables ¿qué utilidad pueden tener entonces?

Algunos como Paul Thagard (2012; pp. 10-14)) se han preocupado por los peligros que los experimentos mentales encierran al ser parte constitutiva de la metodología en la ciencia. Thagard considera que los experimentos mentales tienen un rol legítimo en

la generación, clarificación de hipótesis e identificación de problemas en la ciencia, pero afirma que cuando se pretende extender esta función es cuando los peligros aumentan, como por ejemplo suponer que los experimentos mentales pueden proveer evidencia que por sí sola sustente o justifique la aceptación o rechazo de creencias.

Para Thagard, aceptar hipótesis únicamente sobre la base de pensar acerca de ellas constituye un exceso de confianza con muchas consecuencias negativas, como la adquisición de creencias falsas y la obstrucción de otras vías de investigación. Su postura respecto a los experimentos mentales es escéptica cuando éstos son usados como evidencia para influir en la aceptación o rechazo de hipótesis en la ciencia; los experimentos mentales no constituyen por sí solos evidencia fiable, sino que éstos tienen que ser evaluados con respecto a la evidencia empírica que es la que nos conecta con el mundo, de lo contrario todo se reduce a meras ficciones.

De acuerdo con Thagard (Cfr. Thagard, 2012; pp. 10-14), algunos peligros a los que los experimentos mentales conducen cuando los usamos más allá de su rol generativo de hipótesis y para mostrar posibles inconsistencias en otras teorías son:

1. A menudo llevan a falsas conclusiones.
2. Describen situaciones confusas y ambiguas que permite asumir que cualquier cosa imaginada es de hecho posible.
3. La justificación ignora la evidencia relevante.
4. Caen en un razonamiento circular que inicia con una hipótesis y luego se genera pseudo-evidencia que sostiene a esa hipótesis.
5. Pueden obstaculizar vías de investigación.
6. Pérdida de tiempo.
7. Pueden desacreditar a la filosofía.

En esta misma línea podemos encontrar a Pierre Duhem (1913); su preocupación gira en torno a la lejanía que algunos experimentos mentales guardan con el mundo real y afirma que los experimentos mentales son quimeras porque no se realizan y porque estrictamente no tienen la posibilidad de ser realizados. Para Duhem los experimentos mentales no constituyen una metodología fiable, la ciencia debe evitar hacer uso de ellos y enfocarse únicamente en la experimentación física.

Sin embargo, estos peligros que señalan Thagard y Duhem no son exclusivos de los experimentos mentales, sino que son peligros que se presentan a la investigación científica en general, éstos no son mayores que los peligros que encierra la experimentación física. Así mismo, es plausible sugerir que para que un experimento

mental pueda considerarse significativo y científicamente relevante, una condición fundamental debería ser que éste se encuentre inmerso en una teoría específica, basarse en los datos que la teoría arroja y estar en conexión con la aparición de nueva evidencia o reinterpretación de datos, de lo contrario los experimentos mentales pueden tornarse en razonamientos carentes de sustento teórico, interés y relevancia, cuyo resultado es como menciona Thagard, llevarnos a la adquisición de falsas creencias.

La característica metodológica de los experimentos mentales es uno de los elementos más importantes, ya que los experimentos mentales constituyen antes que nada herramientas metodológicas, herramientas en y para la investigación científica y filosófica. Éstos pueden ser concebidos, a su vez, como herramientas *a priori*, que poseen un carácter contrafáctico, una dimensión heurística, entre otras cualidades, pero antes de todas estas cualidades y características, los experimentos mentales constituyen herramientas metodológicas para la investigación. Por ello, una de las investigaciones más relevantes sobre los experimentos mentales es la que se enfoca en su fiabilidad, es decir, en las condiciones que deben cumplir para constituir una metodología fiable y por tanto útil y científicamente relevante.

2. Sobre la utilidad de los experimentos mentales.

¿Constituyen los experimentos mentales herramientas metodológicas útiles en la ciencia? Como se mencionó en la sección anterior, hay quienes consideran, como Thagard, que los experimentos mentales encierran demasiados peligros y que cuando éstos se utilizan más allá de un rol generativo de hipótesis o para mostrar posibles inconsistencias en las teorías, pueden llevarnos a adquirir falsas creencias; por lo que en estos casos los experimentos mentales no pueden constituir herramientas útiles en la ciencia. Algunos otros, como Duhem, consideran que la ciencia debe enfocarse en la experimentación física y evitar la experimentación mental; por lo que nuevamente los experimentos mentales no se consideran herramientas útiles en la ciencia.

Sin embargo, una gran cantidad de científicos y filósofos, entre otros, consideran todo lo contrario y ven en los experimentos mentales una metodología que ha resultado útil y en ocasiones esencial para el desarrollo y construcción de la ciencia. De esta manera, la discusión sobre la utilidad metodológica de los experimentos mentales puede dividirse en dos grandes grupos, aquellos que están a favor y aquellos que

están en contra. La siguiente sección tiene como finalidad resumir de manera general las ideas involucradas en esta discusión.

2.1 Apologetas y detractores.

Quienes consideran que los experimentos mentales no constituyen una herramienta metodológica útil en la ciencia tienen varias ideas en mente. Por un lado consideran que un experimento mental no debe sustituir jamás a un experimento físico y que la ciencia debe evitar usarlos y enfocarse únicamente en la experimentación física y realizable. Por otro lado, señalan que al ser una metodología *a priori* los experimentos mentales son imposibles de ser llevados a cabo físicamente y que en estos casos la única manera de resolverlos es a través de nuestras intuiciones, lo cual es inaceptable ya que nos lleva al problema de la confiabilidad de nuestras intuiciones⁴; en este caso se considera que cuando un experimento mental es demasiado lejano de una situación real, es decir, que no es corroborable a corto plazo, éste deja de ser significativo.

Como se mencionó en el capítulo anterior, la preocupación de Pierre Duhem (1913) gira en torno a estas cuestiones. Para Duhem los experimentos mentales son ficciones que estrictamente no tienen la posibilidad de ser realizados, y por esta razón la ciencia debe evitar usarlos y enfocarse únicamente en la experimentación física realizable.

Mais il y a pis. Bien souvent, l'expérience fictive qu'on invoque est non seulement irréalisée, mais irréalisable; elle suppose l'existence de corps qu'on ne rencontre pas dans la nature, de propriétés physiques qui n'ont été observées [...] (Duhem, 1913; p. 163)⁵

Il nous suffirait de feuilleter les traités et les manuels de Physique pour y reveler une foule d'expériences fictives; nous y trouverions à foison des exemples des diverses formes que peut revêtir une telle expérience, depuis

⁴ El problema de la confiabilidad de las intuiciones supone que no existe un criterio de demarcación universal para poder decir cuáles intuiciones son correctas y cuáles no; es decir no existe un acuerdo universal para saber qué o cual intuición emplear en los experimentos mentales. Algunas propuestas relacionadas son la de James Robert Brown (1991), George Bealer (1998) y Stephen Stich, entre otros.

⁵ Traducción mía: Pero hay cosas peores. Muy a menudo el experimento ficticio (mental) invocado no solamente no es realizado, sino que es irrealizable; presupone la existencia de cuerpos que no se encuentran en la naturaleza, con propiedades físicas que jamás han sido observadas [...]

l'expérience simplement irréalisée jusqu'à l'expérience absurde. Ne nous attardons pas à cette fastidieuse besogne. (Duhem, 1913; p. 164)⁶

Sin embargo, un gran número de filósofos, físicos, matemáticos, psicólogos, sociólogos, etc., no sólo consideran que los experimentos mentales son una herramienta metodológica útil en la ciencia, sino que éstos desempeñan un papel fundamental en la investigación. Ya que a veces no es posible investigar ciertas áreas a través de métodos experimentales estrictamente físicos, por lo que los experimentos mentales constituyen una herramienta muy útil para investigar en estos casos.

Los defensores consideran además, que la cualidad *a priori* es un elemento característico de todo experimento mental y señalan que el hecho de que un experimento mental no sea, en principio, realizable físicamente no debe ser una preocupación. A diferencia del primer grupo, estos autores consideran que en algunos casos los experimentos mentales pueden ser llevados a cabo físicamente, siempre y cuando el experimento mental no viole las leyes de la naturaleza –aquellas aceptadas en el mundo actual. De esta manera, reconocen la utilidad de los experimentos mentales y no tienen problema en aceptar que éstos pueden, eventualmente y en ciertos casos, ser realizados de manera física.

Finalmente, los defensores apelan a la historia de la ciencia y señalan que ésta ha avanzado haciendo uso de un gran número de experimentos mentales lo cual, de cierta manera, legitima su uso y muestra su gran utilidad. Entre los principales apologetas podemos encontrar a Alexander Koyré, quien consideró que la buena física se hace *a priori* y estableció para Galileo la gloria y el mérito por haber sabido cómo prescindir de los experimentos físicos (Cfr. Koyré, 1968; p. 68). Para Koyré los experimentos mentales no sólo constituyen una metodología útil en la ciencia, sino que demostraron un aspecto importante de la naturaleza del conocimiento, su dimensión sintética *a priori*. Considera además, que su función de idealización es esencial en el pensamiento científico, ya que ésta hace posible la construcción de situaciones hipotéticas o ideales, a través de las cuales exploramos algún aspecto en particular, y esto sólo puede ser llevado a cabo en el laboratorio de la mente.

Pero ¿qué es lo que está de fondo entre estas posturas contrapuestas? Desde mi punto de vista, es plausible sugerir que cada una de las dos posturas en esta discusión, a saber, la que está a favor de la utilidad de los experimentos mentales y la

⁶ Traducción mía: Si volvemos a las páginas de los tratados y manuales de Física podemos recoger un buen número de experimentos ficticios (mentales); encontraríamos allí abundantes ilustraciones de las variadas formas que tal experimento puede asumir, desde el experimento meramente no llevado a cabo hasta el experimento absurdo. No derrochemos tiempo en tan fastidiosa tarea.

que está en contra, está asumiendo o partiendo de un tipo de posibilidad distinta. La siguiente sección presenta una consideración al respecto, la cual tiene como objetivo clarificar la naturaleza de esta discusión y presentar una perspectiva distinta para abordarla.

2.2 Una consideración desde la modalidad.

En esta sección se presenta de manera general la noción de modalidad y la manera en que ésta puede entenderse a partir de varios enfoques, intentando con ello arrojar luz sobre la discusión de la utilidad de los experimentos mentales. La hipótesis es que dependiendo de qué tipo de posibilidad se priorice se tiende a aceptar la utilidad de los experimentos mentales o a rechazarla.

Las definiciones que presento son las que desarrolla Anand Vaidya⁷ y la misma autora acepta que intentar definir la modalidad es una tarea por demás problemática y controversial. También cabe mencionar que su propuesta no es estrictamente lógica o formal como la *kripkeana*, sino que se trata de una visión que intenta definir las estructuras modales de forma general.

De esta manera, la modalidad se puede entender como los modos en que son las cosas o la manera en que nos expresamos acerca de ellas. Las dos formas básicas en las que nos podemos expresar son *la necesidad y la posibilidad* y éstas se pueden entender, a su vez, desde distintos enfoques.

Así tenemos por ejemplo: *la necesidad y la posibilidad epistémica, la necesidad y posibilidad lógica, la necesidad y posibilidad conceptual, la necesidad y posibilidad metafísica y la necesidad y posibilidad física o nomológica*. Particularmente me interesa señalar la posibilidad física y metafísica, ya que son éstas las que me son útiles para intentar arrojar luz sobre la discusión sobre la utilidad de los experimentos mentales. Si bien puede considerarse que las definiciones que presentaré son discutibles, considero que éstas rescatan la idea principal de la estructura modal correspondiente y esto es suficiente para el propósito antes mencionado.

De esta manera, es posible entender la necesidad y posibilidad epistémica de la siguiente forma:

⁷ En *The epistemology of modality*, Stanford Encyclopedia of Philosophy, 2007.

P es *epistémicamente necesario* para un sujeto S si y sólo si $\sim P$ se descarta por lo que S sabe.

Ejemplo: Dado todo lo que Juan sabe, es necesario que hoy haya examen.

P es *epistémicamente posible* para un sujeto S si y sólo si P no se descarta por lo que S sabe.

Ejemplo: Dado todo lo que Juan sabe, es posible que hoy haya examen.

La necesidad y posibilidad epistémica puede interpretarse también a nivel de comunidad epistémica, y así tenemos en términos más generales que P es *epistémicamente necesario* para una comunidad científica C si y sólo si $\sim P$ se descarta por lo que la comunidad científica C sabe. Y que P es *epistémicamente posible* para una comunidad científica C si y sólo si P no se descarta por lo que la comunidad científica C sabe.

La necesidad y posibilidad lógica como:

P es *lógicamente necesario* en un sistema S si y sólo si P es un axioma de S o derivable de los axiomas de S a través de las reglas de S .

Ejemplo: Dada la lógica proposicional, $P \rightarrow P$ es necesario.

P es *lógicamente posible* en un sistema S si y sólo si P es consistente con los axiomas del sistema S .

Ejemplo: Dada la lógica proposicional, la proposición "Habrà examen hoy" es lógicamente posible.

La necesidad y posibilidad conceptual como:

P es *conceptualmente necesario* si y sólo si $\sim P$ se descarta por el conjunto de todas las verdades conceptuales. Los enunciados son verdaderos en virtud de sus significados; se trata de una verdad semántica.

Ejemplo: Es conceptualmente necesario que los triángulos tengan tres lados.

P es *conceptualmente posible* si y sólo si P no se descarta por el conjunto de todas las verdades conceptuales.

Ejemplo: Es conceptualmente posible que la tierra sea plana.

La necesidad y posibilidad metafísica como:

P es *metafísicamente necesario* si y sólo si P es verdadero en todos los mundos metafísicamente posibles.

Ejemplo: Es metafísicamente necesario que Aristóteles el filósofo griego sea humano.

Y P es *metafísicamente posible* si y sólo si P es verdadero en algún mundo metafísicamente posible.

Ejemplo: Es metafísicamente posible que alguna partícula física se mueva más rápido que la velocidad de la luz.

La necesidad y posibilidad física o nomológica como:

P es *físicamente necesario* con respecto a las leyes físicas L si y sólo si P está lógicamente implicado por L.

Ejemplo: Dadas las leyes actuales de la física, es físicamente necesario que ninguna partícula viaje más rápido que la velocidad de la luz.

Y P es *físicamente posible* con respecto a las leyes físicas L si y sólo si P es lógicamente consistente con L.

Ejemplo: Dadas las leyes actuales de la física, es físicamente posible para un tren viajar a 30 k/hr.

De esta manera, la primera postura que afirma que los experimentos mentales no constituyen una herramienta metodológica útil en la ciencia parece que está priorizando de alguna forma la *posibilidad física* ante todo, ya que argumenta, en primer lugar, que los experimentos mentales o gran parte de ellos son imposibles de ser llevados a cabo físicamente y, en segundo lugar, que los experimentos mentales no deben sustituir jamás a los experimentos físicos o realizables. Priorizar la *posibilidad física o nomológica* explica el rechazo de la utilidad de los experimentos mentales en la ciencia, ya que éstos no siempre apelan a una *posibilidad física* estrictamente, sino que una gran variedad de experimentos mentales, en distintas áreas del conocimiento, apelan a la *posibilidad metafísica o conceptual*.

En filosofía de la mente, por ejemplo, la *posibilidad metafísica* es por excelencia el tipo de posibilidad a la que los experimentos mentales apelan. Así tenemos experimentos mentales como “Mary la científica del color” de Jackson⁸, “la tierra gemela” de

⁸ Jackson, Frank (1982), “Epiphenomenal Qualia”, *Philosophical Quarterly*, No. 32, pp. 127-136.

Putnam⁹, “la habitación china” de Searle¹⁰, “el hombre del pantano” de Davidson¹¹, entre otros. Estos experimentos mentales priorizan la *posibilidad metafísica*, es decir, exploran situaciones hipotéticas donde P no necesita ser verdadero en el mundo actual sino en algún mundo metafísicamente posible.

Por otro lado, la segunda postura que afirma que los experimentos mentales sí constituyen una herramienta metodológica útil en la ciencia, parece que está considerando la utilidad de la *posibilidad metafísica*, ya que afirma que existen ciertas áreas de la realidad que no se pueden investigar a través de medios estrictamente físicos y que en estos casos la experimentación mental es de gran ayuda. De igual manera, apelar o priorizar la *posibilidad metafísica* no excluye que los experimentos mentales apelen a otros tipos de posibilidad. Un experimento mental en física por ejemplo, puede apelar a la *posibilidad metafísica* por la imposibilidad de ser realizado físicamente en un determinado momento, ya sea por causas de la tecnología actual o por causas de imposibilidad implícita, y apelar al mismo tiempo a la *posibilidad física o nomológica* - como el caso del experimento mental de los gemelos de Einstein¹².

La mayoría de los experimentos mentales en física apelan o priorizan la *posibilidad física o nomológica*, como el experimento mental de Galileo y las balas de cañón¹³. En cambio, los experimentos mentales en filosofía de la mente, en epistemología o en ética apelan más bien a la *posibilidad metafísica, epistémica o conceptual*. Así, en filosofía de la mente será metafísicamente posible, por ejemplo, que exista una criatura humanoide carente de conciencia; en otras palabras, que exista un zombi es metafísicamente posible si y sólo si que exista un zombi es verdadero en algún mundo metafísicamente posible.

Como ya se mencionó, priorizar un tipo de posibilidad sobre otra no implica que los experimentos mentales no involucren otros tipos de posibilidad, como la *conceptual*, la *lógica* o la *epistémica*, entre otras. Engel (2011) por ejemplo, sostiene que los experimentos mentales implican *posibilidades conceptuales* y Williamson (2007) que

⁹ Putnam, Hilary (1973), “Meaning and Reference”, The Journal of Philosophy, Vol. 70, No. 19, pp. 699-711.

¹⁰ Searle, John (1980), “Minds, brains and programs”, Behavioral and Brain Sciences, No. 3, pp. 417-457.

¹¹ Davidson, Donald (1987), “Knowing One’s Own Mind”, Proceedings and Addresses of the American Philosophical Association, Vol. 60, No. 3, pp. 441-458.

¹² El experimento mental de los gemelos fue propuesto originalmente como el experimento mental de los relojes y éste tiene su origen en el trabajo de Einstein de 1905 donde sienta las bases de la Teoría de la Relatividad Especial. La formulación del experimento mental de los relojes se puede consultar en: Einstein, Albert (1998) “On the Electrodynamics of Moving Bodies” en Stachel, J. (ed.), Einstein’s Miraculous Year, Princeton University Press, pp.123-160. Mientras que la versión que reemplaza a los relojes por gemelos se puede consultar en: Langevin, Paul (1911), L’Évolution de l’espace et du temps, Scientia 10: 31-54.

¹³ Este ejemplo se desarrollará a detalle en la sección 5.2

refieren o implican *posibilidades metafísicas*. Para Engel la postura de Williamson no parece mostrar que los experimentos mentales apelen únicamente a *posibilidades metafísicas* e intenta mostrar que éstos apelan a *posibilidades conceptuales* también, e incluso que la postura de Williamson es compatible con la suya.

Para Engel los experimentos mentales parecen tener su hogar en la filosofía, la cual es, desde su punto de vista, una disciplina que trabaja esencialmente con conceptos. Por ello, encuentra la propuesta de Williamson muy poco plausible –aunque no errónea- ya que éste afirma que los experimentos mentales tratan con *posibilidades metafísicas*, es decir, con posibilidades acerca de cómo son las cosas, no acerca de la manera en que nosotros pensamos acerca de ellas. Para Engel muchos experimentos mentales no se refieren al mundo, sino a nuestros conceptos acerca de él; los experimentos mentales acerca de la identidad personal por ejemplo, analizan nuestro concepto de identidad personal no a la identidad personal como algo en sí mismo. El experimento mental del auditorio de Strawson¹⁴ trata con nuestro concepto de un mundo objetivo y así, podemos afirmar que muchos experimentos mentales exploran los límites y extensión de nuestros conceptos, ya sean ordinarios o filosóficos. Sin embargo, aunque Engel se posiciona desde la tradición analítica para decir que la filosofía trata con conceptos y no con las cosas en sí, éste afirma que esto no significa que los experimentos mentales no se refieran a las cosas del mundo o a posibles estados de cosas, sino que la mayor parte del tiempo tenemos acceso a ellas a través de nuestros conceptos.

Thought experiments about persons or about freedom could hardly be about persons or freedom in themselves, they are first and foremost about our concepts of person and of freedom, even though they are supposed to show us something about persons and freedom.

(Engel, 2011; p. 157)

Desde la postura de Engel, los experimentos mentales son una herramienta de crítica, éstos nos ayudan a examinar nuestras intuiciones, a ilustrar los conflictos, analizar nuestros conceptos y a ofrecer hipótesis para argumentar acerca de un estado de cosas. Éstos no sólo tratan con *posibilidades metafísicas* como supone Williamson, sino también con *posibilidades conceptuales*.

¹⁴ Strawson, P. F. (1959), *Individuals . An Essay in Descriptive Metaphysics*, Routledge, London and New York, pp. 64-86.

So it does not seem to me that Williamson has shown that all thought experiments deal with metaphysical and not conceptual possibilities, although I agree that in philosophy we have both.

(Engel, 2011; p. 162)

De esta manera, para Engel el razonamiento filosófico puede ser conceptual y metafísico. Los experimentos mentales pueden apelar a distintos tipos de posibilidad; así tenemos experimentos mentales que apelan o priorizan la *posibilidad metafísica* y experimentos mentales que apelan o priorizan la *posibilidad conceptual*.

Como puede observarse, la discusión entre apologetas y detractores puede explicarse desde la perspectiva de la modalidad. Cuando se prioriza la *posibilidad física* se tiende a rechazar la utilidad de los experimentos mentales; en cambio cuando se apela o prioriza la *posibilidad metafísica* se tiende a aceptar dicha utilidad. Considero que ambos tipos de posibilidad son importantes; la *posibilidad metafísica* es esencial para el análisis y la reflexión, ya que gran parte de nuestras teorías científicas y filosóficas se basan en experimentos mentales que apelan a *posibilidades metafísicas*; mientras que la *posibilidad física* permite corroborar empíricamente lo que el experimento mental ilustra. Desde mi punto de vista los dos tipos de posibilidad, lejos de excluirse, se complementan, pues la experimentación mental es, en ocasiones, la experimentación previa a la experimentación física, siendo ésta última la ejecución empírica de lo que ya se ejecutó en la mente o el pensamiento. De esta forma, es posible decir, que los experimentos mentales constituyen una herramienta metodológica útil en la ciencia; sin embargo éstos no son la única herramienta, sino que forman parte de la caja de herramientas que el científico y el filósofo utilizan. Un experimento mental es en ocasiones una forma de experimentación previa y a veces la única forma de experimentación a la que podemos aspirar. Esto no quiere decir que un experimento mental no pueda realizarse nunca, como ya se mencionó, es posible llevar a cabo algunos experimentos mentales cuando estos son teóricamente y tecnológicamente posibles.

La discusión sobre la utilidad de los experimentos mentales es por si misma fructífera, ya que si se concede que los experimentos mentales constituyen una herramienta metodológica útil en la ciencia, emergen nuevas interrogantes como:

¿Qué tipo de herramienta?

¿Para qué son útiles en la ciencia?

Mi propuesta respecto a los dos puntos anteriores es la siguiente: *los experimentos mentales constituyen una herramienta metodológica que el científico o el filósofo utiliza para explorar, fortalecer o debilitar hipótesis, así como para argumentar en la ciencia, esto es, para ofrecer razones a favor o en contra de una determinada teoría.*

Pero si a través de los experimentos mentales se argumenta, cabe preguntarse entonces qué tipo de estructura argumentativa involucran los experimentos mentales. La sección siguiente presenta una de las principales discusiones sobre los experimentos mentales, a saber, si éstos involucran una estructura argumentativa y de ser así de qué tipo.

3. La estructura argumentativa de los experimentos mentales.

Desde la visión actual es ampliamente aceptado que ningún experimento en general es conclusivo, es decir, que ningún experimento demuestra o refuta una hipótesis o teoría de manera concluyente y esto vale también para los experimentos mentales. Por ello, como ya dijo Duhem¹⁵, lo único que podemos hacer a través de los experimentos en general es ofrecer razones para que nuestras hipótesis o teorías se refuercen o debiliten y es en este sentido que los experimentos mentales constituyen una manera de argumentar en la ciencia. Por lo que cabe preguntarse entonces ¿qué tipo de estructuras argumentativas involucran los experimentos mentales?

Una estructura argumentativa es una estructura constituida por un conjunto de proposiciones o premisas a partir de las cuales se deriva una conclusión; éstas pueden ser deductivas, inductivas, abductivas¹⁶, entre otras. Mientras que a partir de las estructuras deductivas se derivan conclusiones necesarias, a partir de las estructuras inductivas se derivan conclusiones probables y a partir de las estructuras abductivas conclusiones plausibles. La deducción preserva la verdad e implica poder derivar una sola conclusión, la inducción enumerativa postula generalizaciones y la abducción sugiere posibles y plausibles hipótesis explicativas.

De esta manera, algunos autores como Nicholas Rescher (1991), Ian Hacking (1992), James Robert Brown (1991), Andrew D. Irvine (1991), John Forge (1991) y John Norton (1991), entre otros, aceptan que los experimentos mentales tienen una

¹⁵ Duhem, Pierre (1914), *La théorie physique: son objet-sa structure*, segunda parte, cap. vi, *La théorie physique et l'expérience*, pp. 146-176.

¹⁶ Esto se desarrolla a detalle en el punto 4.1

estructura argumentativa y que éstos constituyen una forma de argumentar, ya sea a favor o en contra de una determinada hipótesis o teoría.

A “thought experiment” is an attempt to draw instruction from a process of hypothetical reasoning that proceeds by eliciting the consequences of an hypothesis which, for aught that one actually knows to the contrary, may well be false. It consists in reasoning from a supposition that is not accepted in the interests of making a point or resolving a conclusion.

(Rescher, 1991; p. 31)

The reason that people wrestle with thought experiments, use them for exposition and put-down arguments, is that they can reveal tensions between one vision of the world and another. They can dislodge a person from a certain way of describing the world. They can replace one picture by another. That is their job, their once and future job.

(Hacking, 1992; p. 307)

As its name suggest, the destructive thought experiment is essentially an argument directed against a theory. It is a picturesque *reduction ad absurdum*: it destroys or at least present serious problems for a theory, usually by pointing out a shortcoming in its general framework. Such a problem may be anything from a minor tension with other well-entrenched theories to an outright contradiction within the theory itself.

(Brown, 1991; p. 34)

[...] thought experiments, in the first instance, are simply arguments concerning particular events or states of affairs of a hypothetical (and often counterfactual) nature which lead to conclusions about the nature of the world around us. After all, thought experiments do not purport to introduce any new empirical data and, since they do purport to yield some new information about the world, the only other source of obtaining such new information must result from the reconsideration of previous data by way of argument. This much, at least, should be uncontroversial.

(Irvine (1991; p. 150)

Como puede observarse, todos ellos destacan la función argumentativa de los experimentos mentales, aunque la manera en que formulan esta idea es distinta. La mayoría la expresa diciendo que los experimentos mentales *son argumentos*, yo prefiero decir que éstos *involucran estructuras argumentativas*, ya que pensar en estructuras argumentativas evita compromisos reduccionistas, sin embargo, me apegaré a la forma en que cada autor prefiere expresarse.

Engel (2011) por ejemplo se expresa con la primera formulación y dice que los experimentos mentales son argumentos, afirma que cuando los filósofos usan experimentos mentales lo hacen para argumentar a favor o en contra de alguna noción o concepto determinado, por ejemplo como cuando discuten acerca de las diferentes nociones del conocimiento. Para Engel los experimentos mentales son constructos que pretenden llevarnos de premisas verdaderas a conclusiones verdaderas, son hipótesis que proponemos y desarrollamos en conexión con otros supuestos, en algunos casos las deliberaciones son convincentes, en otros casos no.

Para Norton (1991, 2004) los experimentos mentales son argumentos de tipo deductivo -o utilizando la segunda formulación, que éstos involucran principalmente estructuras argumentativas de tipo deductivo-, es decir, estructuras que constan de ciertas premisas y una conclusión. Norton considera que través de los experimentos mentales explicamos y obtenemos conocimiento del mundo y que éstos pueden ser fiables si son gobernados por algún tipo de lógica como la deductiva. También asegura que los experimentos mentales no trascienden el empirismo ya que nos hablan del mundo, por lo que éstos hacen uso de lo que ya sabemos, explícita o tácitamente, y así transforman el conocimiento existente en una argumentación pictórica o en una narrativa.

[...] thought experiments in science are merely picturesque arguments. Their epistemic reach can always be replicated by an argument and this is best explained by their merely arguments.

(Norton, 2004; p. 64)

Norton considera que los experimentos mentales tienen las siguientes características:

1. Posit hypothetical or counterfactual states of affairs, and
2. Invoke particulars irrelevant to the generality of the conclusion.

(Norton, 1991; p. 85)

Así, un experimento mental puede reconstruirse como un argumento deductivo, haciendo evidente que consta de ciertas premisas y que lleva a una conclusión.

Cabe mencionar que para Norton, la explicación de por qué los experimentos mentales involucran estructuras argumentativas se basa en la idea de que el pensamiento puro no puede producir conocimiento, a excepción de las verdades lógicas, sino que todo lo que el pensamiento puro puede hacer es transformar lo que ya sabemos. La idea es que si los experimentos mentales nos dan conocimiento del mundo es porque sus premisas contienen información empírica y es por esta razón que pueden ser reconstruidos como argumentos deductivos. Para Norton, los experimentos mentales tienen el mismo poder epistémico que los argumentos y considera que a través de ellos justificamos y obtenemos conocimiento.

Thought experiments in physics provide or purport to provide us information about the physical world. Since they are *thought* experiments rather than physical experiments, this information does not come from the reporting of new empirical data. Thus there is only one non-controversial source from which this information can come: it is elicited from information we already have by an identifiable argument, although that argument might not be laid out in detail in the statement of the thought experiment. The alternative to this view is to suppose that thought experiments provide some new and even mysterious route to knowledge of the physical world.

(Norton, 1991; p. 129)

John Forge se apega a esta definición deductivista (Cfr. Forge, 1991; p. 210) y considera igualmente que los experimentos mentales se sustentan en estructuras deductivas. Sin embargo, autores como Michael Bishop (1999) consideran que los experimentos mentales no se sustentan en estructuras argumentativas de tipo deductivo¹⁷. Bishop ataca la caracterización que Norton (1991) presenta y afirma que la propuesta deductivista colapsa, ya que algunos experimentos mentales no se adaptan a la estructura lógica de premisas-conclusión; y para mostrar esto recurre al experimento mental del reloj en la caja que Einstein presentó durante la conferencia Solvay de 1930 sobre magnetismo y el cual construyó como un contraejemplo al

¹⁷ Yo concuerdo con Bishop en que la propuesta deductivista de Norton colapsa en ciertos casos, pero las reconstrucciones deductivas son un tipo de reconstrucción, no la única. Por lo que habrá casos en los que la estructura deductiva sea exitosa, mientras que en otros casos no.

principio de incertidumbre de Heisenberg¹⁸. A continuación una cita que ilustra de manera breve en qué consiste básicamente este experimento mental:

Suppose we have a box full of protons that has, on one of its walls, a shutter that is controlled by a clock. Weigh the box. Now set up the shutter mechanism so that it opens for a brief interval at which time a single photon escapes. Weigh the box again. The change in the weight of the box gives us the weight of the photon, which gives us its mass. And using Einstein's famous equation, $E=mc^2$, we can determine the photon's energy. In principle, therefore, we can measure the photon's energy and its time of passage to any arbitrary degree of accuracy. So on the basis of this thought experiment, Einstein concluded that Heisenberg's uncertainty principle is false.

(Bishop, 1999; p. 536)

De acuerdo con Bishop, el propósito de Einstein era mostrar que existía un experimento a través del cual se podía medir la velocidad y posición exacta de una partícula. En cambio, Bohr defendía que un experimento podía medir, con cierto grado de exactitud, la velocidad o la posición de una partícula pero no ambas. Para demostrar esto, Bohr se enfocó en los instrumentos prácticos y procedimientos que se tendrían que utilizar para medir la energía del fotón en un tiempo particular y mostró que pesar la caja requeriría que el aparato del reloj en la caja se moviera en un campo gravitacional, lo cual demuestra que hay un límite fundamental en la precisión con el que el aparato del reloj en la caja puede medir el peso del fotón (y por lo tanto su energía) en un tiempo particular. De esta manera, reivindicó el principio de incertidumbre de Heisenberg y ante tal explicación Einstein abandonó el intento de socavar el principio de incertidumbre.

Para Bishop el experimento mental del reloj en la caja no puede ser identificado con un argumento de tipo deductivo, considera que Bohr y Einstein están analizando un mismo experimento mental pero que ellos están proponiendo dos argumentos o estructuras deductivas distintas, por lo tanto, el experimento mental del „reloj en la caja“ no puede ser el argumento o la estructura argumentativa deductiva. Los argumentos o las estructuras argumentativas son las que Bohr y Einstein proponen y las cuales llevan a conclusiones distintas y contrarias.

¹⁸ En 1927, Werner Heisenberg introdujo el principio de incertidumbre que lleva su nombre. Este principio afirma que no es posible conocer la velocidad y la posición de una partícula subatómica al mismo tiempo, hay un límite irreducible cuando se intenta medir con precisión la conjugación de un par de variables (variables como la posición-momento y el tiempo-energía). Esta relación se describe con la siguiente ecuación: $\Delta p \times \Delta q > h$. Donde Δp y Δq permanecen representan la incertidumbre de las respectivas variables y h es una constante (la constante de Planck dividida por 2π).

Para Bishop la postura que sostiene que los experimentos mentales son argumentos - o que éstos se sustentan en estructuras argumentativas- está comprometida a la siguiente tesis.

(A) Dos casos de un experimento mental son casos del mismo tipo de experimento mental si y sólo si son casos del mismo tipo de argumento deductivo –o estructura deductiva.

El problema que Bishop ve es que en el experimento mental del reloj en la caja (A) es falsa. Se trata de un experimento mental idéntico en tipo que lleva a dos conclusiones distintas y contrarias. El experimento mental del reloj en la caja es un experimento mental del mismo tipo pero los argumentos de Bohr y Einstein son casos diferentes de argumentos, es decir, los argumentos o estructuras argumentativas propuestas por Bohr y Einstein no son idénticos en tipo cuando deberían serlo ya que se trata del mismo tipo de experimento mental. Por lo tanto, la postura que afirma que los experimentos mentales son argumentos deductivos –o que se basan en estructuras deductivas- es falsa, ya que (A) no se cumple.

Bishop considera que la mejor vía para eludir esta objeción, sería sostener que el experimento mental de Einstein no es idéntico en tipo al de Bohr, es decir, que no se trata en el fondo del mismo experimento mental a pesar de sus similitudes. Sin embargo, encuentra esta vía muy poco plausible, ya que si Bohr no hubiera replicado el experimento mental del „reloj en la caja“, entonces no hubiera sido posible ejecutar ese mismo experimento mental apropiadamente y Einstein podría haberlo acusado de cambiar de tema. En cambio, la actitud de Einstein muestra que se trataba en efecto del mismo experimento mental, de lo contrario qué razones habría tenido para abandonar la contrarréplica de Bohr. La única manera de dar sentido a esto es suponer que Bohr repitió el experimento mental del „reloj en la caja“ original. Por lo tanto, se trataba en esencia del mismo experimento mental y no de dos experimentos mentales distintos.

Para Bishop el experimento mental del reloj en la caja muestra que los experimentos mentales no pueden ser argumentos de tipo deductivo o al menos que éstos no se basan necesariamente en estructuras deductivas, ya que a partir de un mismo experimento mental se derivaron dos conclusiones distintas y contrarias. Los argumentos o estructuras argumentativas son las razones y formulaciones que Bohr y Einstein dieron, no el experimento mental.

Mi postura aquí coincide con la de Bishop en que una reconstrucción deductivista del experimento mental del reloj en la caja falla, ya que a partir del experimento mental se derivan dos conclusiones diferentes y contrarias, pero considero que el experimento mental se puede reconstruir con una estructura argumentativa diferente. Como ya mencioné, mi compromiso no es defender que los experimentos mentales *son argumentos* de algún tipo, sino que éstos involucran *estructuras argumentativas* o que pueden *ser reconstruidos argumentativamente*.

Por otro lado, algunos autores como James Lennox (1991) rechazan que los experimentos mentales se sustenten en argumentos. Para Lennox los experimentos mentales no son argumentos, considera que las reconstrucciones que hace Norton son reducciones o partes de un argumento pero que éstas no son argumentos en sí mismos, sino descripciones hipotéticas que están inmersas en una teoría. Para Lennox un experimento no es un conjunto de proposiciones, sino una actividad humana enfocada en producir interacciones físicas concretas, como observar si los efectos predichos por una teoría se producen.

Para Tamar Szabó Gendler (2004) los experimentos mentales tampoco pueden reducirse a argumentos, ya que envuelven mecanismos cognitivos complejos que son irreducibles a estructuras argumentativas. Para Gendler, el rol que juegan las intuiciones, las emociones y los sentimientos es parte ineludible de la experimentación mental y la manera en que desarrollemos la narrativa influye directamente la conclusión final, esto es, en la aceptación o rechazo de una determinada hipótesis o teoría.

Desde mi punto de vista Lennox y Gendler tienen razón al decir que los experimentos mentales constituyen una actividad humana y que hay elementos que son irreducibles, pero esto no significa que no podamos realizar una reconstrucción lógica a partir de aquellos elementos que sí se pueden utilizar. Considero que si bien, no es posible reducir los experimentos mentales a meros argumentos por las razones antes esbozadas, es posible en cambio, intentar reconstruirlos argumentativamente.

Como puede observarse, hay quienes consideran que los experimentos mentales son argumentos o que es posible reconstruirlos argumentativamente y quienes consideran que no. La discusión entre Norton y Bishop rescata a grandes rasgos esta diferencia de posturas, ya que aquellos que aceptan que los experimentos mentales son argumentos o que involucran una estructura argumentativa lo hacen desde una posición muy cercana, si no es que idéntica a la de Norton; mientras que Bishop es quien lanza un ataque directo a esta propuesta específica, por considerar que la

caracterización deductivista que propone Norton colapsa, ya que un mismo experimento mental, constituido de ciertas premisas, puede dar lugar a dos conclusiones distintas, lo cual entra en contradicción con la idea de un argumento deductivo, donde hay una única conclusión derivada con necesidad de las premisas, esto es, donde la relación entre premisas y conclusión se clausura bajo el concepto de consecuencia lógica. Sin embargo, desde una perspectiva abductivista los experimentos mentales no estarían obligados a derivar con necesidad una única conclusión o explicación, sino un conjunto de explicaciones posibles y plausibles; lo que explicaría por qué un mismo tipo de experimento mental puede dar lugar a dos o más conclusiones o explicaciones distintas y el hecho de que ciertos experimentos mentales hayan sido aceptados en un determinado momento en el desarrollo de la ciencia y rechazados en otro periodo. En la siguiente sección presentaré en qué consiste el razonamiento abductivo y en qué sentido los experimentos mentales pueden ser concebidos como argumentos abductivos o de plausibilidad.

4. Razonamiento abductivo.

El objetivo de este capítulo es presentar en qué consiste el razonamiento abductivo y en qué se diferencia de otros tipos de razonamientos como el deductivo, el inductivo y la inferencia a la mejor explicación. Las propuestas en las que me baso para desarrollar este capítulo son las de Charles Sanders Peirce (1958) y la de la Dra. Atocha Aliseda Llera (2006).

Charles Sanders Peirce (1839-1914) fue el primer filósofo que profundizó en el estudio del razonamiento abductivo como tal y el primero en dar a la abducción una forma lógica. Para Peirce la abducción es el tipo de razonamiento que parte de una observación hacia sus posibles hipótesis explicativas.

Long before I first classed abduction as an inference it was recognized by logicians that the operation of adopting an explanatory hypothesis -- which is just what abduction is -- was subject to certain conditions. Namely, the hypothesis cannot be admitted, even as a hypothesis, unless it be supposed that it would account for the facts or some of them. The form of inference, therefore, is this:

The surprising fact, C, is observed;

But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true.

(Peirce, 1958; CP 5.189)

Para Peirce la abducción engloba tres aspectos importantes: debe ser explicativa, corroborable y económica. Para ser explicativa, la abducción debe poder dar cuenta del hecho observado, es decir, debe explicar los hechos. La corroboración se refiere a que la hipótesis explicativa debe ser capaz de poder ser corroborada, aunque ésta no se limita a la corroboración empírica o a la verificación en términos de Peirce.

[...] if pragmatism is the doctrine that every conception is a conception of conceivable practical effects, it makes conception reach far beyond the practical. It allows for any flight of imagination, provided this imagination ultimately alights upon a possible practical effect; and thus many hypothesis may seem at first glance to be excluded by the pragmatism maxim that are really not excluded.

[Peirce, 1958; CP, 5. 196]

De esta manera, la corroboración se refiere también al proceso de concebir posibles efectos, ya que calcular los efectos que una hipótesis explicativa produce involucra en sí mismo experimentación y corroboración. Para Aliseda, lo anterior sugiere un espacio que da lugar a experimentos mentales y a otras manifestaciones en el ámbito de las ideas (Cfr. Aliseda, 2006; p. 176). Así, la corroboración puede ser entendida pues desde un aspecto práctico o desde un aspecto teórico.

Finalmente, el criterio de economía permite afrontar el problema de tener un conjunto innumerable de posibles hipótesis explicativas y la necesidad de elegir la hipótesis que mejor explique el hecho observado dada la evidencia disponible. (Cfr. Aliseda, 2006; p. 36)

Así, la abducción es concebida como un tipo de razonamiento explicativo, un proceso de razonamiento que utilizamos para explicar un fenómeno o hecho desconocido. *“Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea [...]”* (Cfr. Peirce, 1958; CP 5.171).

Su forma lógica puede expresarse con la siguiente estructura:

C

$[A_1, A_2, \dots A_n]$ $A_1 \rightarrow C$

\therefore Plausiblemente A_1

Donde C es un hecho o fenómeno observado, $[A_1, A_2, \dots A_n]$ es el conjunto de explicaciones generadas y donde A_1 es la explicación seleccionada como la más plausible del conjunto de explicaciones generadas.

El razonamiento abductivo se puede observar en diversas situaciones de la vida diaria o de la práctica científica. Por ejemplo, si nos despertamos y el césped está mojado, podemos explicar esta observación asumiendo que debió haber llovido o que los aspersores de la casa debieron haberse encendido. O si llegamos al salón de clases y no hay nadie, podemos explicar este fenómeno asumiendo que el maestro debió haber cancelado la sesión o que la sesión se debió haber cambiado a otro salón, etc. El diagnóstico médico es por excelencia otro ámbito donde el razonamiento abductivo encuentra lugar; por ejemplo, cuando un médico analiza a un paciente enfermo, éste observa ciertos síntomas en el paciente e hipotetiza acerca de las posibles causas de la enfermedad, y basado en el conocimiento que tiene acerca de las relaciones causales entre las enfermedades genera conjeturas o hipótesis que expliquen los síntomas del paciente. En la práctica científica el razonamiento abductivo se ha discutido en obras como las memorias de Laplace y se ha dicho que el razonamiento de Kepler cuando descubrió que Marte tenía una órbita elíptica fue abductivo también. (Cfr. Hanson, 1972; pp. 72-73). De esta manera, en cualquier situación donde nos enfrentemos ante un hecho u observación desconocida, el razonamiento abductivo nos permite generar distintas explicaciones posibles que den cuenta de tal hecho u observación.

Cabe mencionar que todas las explicaciones abductivas, sean en cualquier ámbito, son falibles, ya que la aparición de nueva información puede invalidar nuestras primeras y tempranas explicaciones. Quizás el césped está mojado porque los niños estuvieron jugando con agua o quizás el salón de clases está vacío porque todos vienen tarde, etc. La nueva evidencia o información puede cambiar nuestras hipótesis o explicaciones anteriores.

4.1 Diferencia entre deducción, inducción y abducción.

Para los fines de este trabajo, se considerará que la abducción es un tipo de razonamiento diferente a la deducción y a la inducción. La abducción no es un caso de inducción, ni de deducción. Si bien, existen propuestas que identifican a la abducción con la inducción, yo consideraré que son tres tipos de razonamiento diferentes, al igual que lo hace Peirce (1958; CP 2. 641, CP 5.171) y Aliseda (2006; p. 33-35).

La deducción es un método de inferencia que deriva conclusiones necesarias, *“Deduction is the only necessary reasoning. It is the reasoning of mathematics.”* (Peirce, 1958; CP, 5. 145)

Deduction proves that something **must** be; Induction shows that something **actually** is operative; Abduction merely suggests that something **may be**.¹⁹

(Peirce, 1958; CP 5.171)

La inducción deriva conclusiones probables y la abducción explicaciones plausibles. La inducción debe ser validada empíricamente con pruebas y experimentos y el razonamiento abductivo ofrece hipótesis que pueden ser refutadas cuando aparece información adicional. Por ello, la abducción es un tipo de razonamiento característico de situaciones donde la información es incompleta (Cfr. Aliseda, 2006; p. 28), aunque cabe mencionar que ésta no constituye una característica exclusiva del razonamiento abductivo, ya que en la inducción enumerativa puede considerarse que la información que con que se cuenta para hacer la inducción es igualmente incompleta.

Como ya se mencionó, los argumentos deductivos son estructuras que constan de ciertas premisas y a partir de las cuales se deriva con necesidad una conclusión. Los argumentos inductivos (inducción enumerativa) por otro lado, son estructuras donde la conclusión sólo se afirma en términos de probabilidad, es decir, donde a partir de las premisas se deriva una conclusión más o menos probable. En cambio, en los argumentos abductivos no se derivan conclusiones necesarias ni probables, sino únicamente plausibles y la aparición de nueva información puede cambiar nuestras primeras y tempranas explicaciones o conclusiones.

¹⁹ Las negritas son del texto original.

De esta manera, el concepto central de la deducción es la **necesidad**; donde la relación entre las premisas y la conclusión se clausura bajo el concepto de consecuencia lógica. Ejemplos²⁰.

[Regla]	1. Todas las canicas de esta caja son blancas	
[Caso]	2. Estas canicas son de esta caja	
[Resultado]	/ Estas canicas son blancas	[Necesario]
[Regla]	1. Todos los planetas giran alrededor del sol	
[Caso]	2. Júpiter es un planeta	
[Resultado]	/ Júpiter gira alrededor del sol	[Necesario]
[Regla]	1. Cada vez que llueve el césped se moja	
[Caso]	2. Llovió	
[Resultado]	/ El césped está mojado	[Necesario]

El concepto central en la inducción (enumerativa) es la **probabilidad**; a partir de la observación de hechos particulares se afirman enunciados generales, esto es, a partir de observaciones de casos particulares se afirma que una generalidad es probable. Ejemplos.

[Caso]	1. Estas canicas son de esta caja	
[Resultado]	2. Estas canicas son blancas	
[Regla]	/ Todos las canicas de esta caja son blancas	[Probable]
[Caso]	1. Marte, Júpiter, Saturno [...] son planetas	
[Resultado]	2. Marte, Júpiter, Saturno [...] giran alrededor del sol	
[Regla]	/ Todos los planetas giran alrededor del sol	[Probable]
[Caso]	1. Llovió	
[Resultado]	2. El césped está mojado	
[Regla]	/ Cada vez que llueve el césped se moja	[Probable]

²⁰ La estructura de estos ejemplos y el contenido del primero, se basan en el libro de la Dra. Atocha Aliseda Llera, *Abductive Reasoning. Logical Investigations into Discovery and Explanation*, Springer, 2006.

En cambio, el concepto central para el razonamiento abductivo es la noción de **plausibilidad**; ante un fenómeno u observación, y a partir de la información teórica que tenemos, es plausible sugerir una determinada hipótesis o explicación que dé cuenta del fenómeno u observación, no en términos de necesidad, no en términos de probabilidad, sino en términos de plausibilidad. Ejemplos.

[Regla] 1. Todas las canicas de esta caja son blancas
 [Resultado] 2. Estas canicas son blancas
 [Caso] / Estas canicas son de esta caja [Plausible]

[Regla] 1. Todos los planetas giran alrededor del sol
 [Resultado] 2. Marte, Júpiter, Saturno [...] giran alrededor del sol
 [Caso] / Marte, Júpiter, Saturno [...] son planetas [Plausible]

[Regla] 1. Cada vez que llueve el césped se moja
 [Resultado] 2. El césped está mojado
 [Caso] / Llovió [Plausible]

La [Regla] se identifica “con el conocimiento previo”, con la “teoría” o con la “información que ya tenemos”; el [Caso] con la “observación realizada” o con el “fenómeno o hecho a explicar”; y el [Resultado] con la “conclusión o explicación generada”.

Por otro lado, el razonamiento abductivo no solamente se puede usar para explicar fenómenos u observaciones realizadas, sino también para intentar explicar lo que pasaría ante un fenómeno determinado. Por ejemplo, la última estructura abductiva se puede reformular de la siguiente manera.

[Teoría] 1. Cada vez que llueve el césped se moja [información previa o teoría]
 [Fenómeno] 2. El césped mojado [fenómeno planteado a explicar, ¿cómo se explicaría que el suelo o césped estuviese mojado?]
 [Explicación] / Llovió [explicación abductiva]

La noción de **plausibilidad** es esencial en el razonamiento abductivo y es el concepto que me interesa rescatar para la caracterización abductivista que estoy proponiendo. Se trata de un concepto pragmático de evaluación cuyos parámetros varían en las comunidades científicas²¹. Esta puede ser entendida como aquello que la comunidad de expertos aceptaría creer por presentarse como una explicación convincente que no genera conflictos, al menos de entrada, con el sistema de creencias establecido. La plausibilidad juega un papel fundamental durante el proceso de generación de posibles explicaciones y en la elección de explicaciones; las explicaciones generadas no sólo son posibles sino también plausibles y la explicación elegida es la explicación considerada la más plausible de todo el conjunto, es decir, la explicación más creíble para un sujeto o para una determinada comunidad científica, aquella explicación que no entre en conflicto o genere inconsistencias con el sistema de creencias compartido.

Así, los valores epistémicos varían en las comunidades científicas, por lo que la mejor explicación no siempre es la misma para un sujeto o una comunidad científica. No obstante, la idea es la misma y elegir la mejor explicación en el razonamiento abductivo significa, elegir la explicación considerada como la más plausible, creíble, confiable o convincente para un sujeto o comunidad, independientemente de que se prioricen valores distintos.

Ahora, de acuerdo con Aliseda (2006) la abducción puede ser usada para referirse a un proceso o a un producto; así podemos hablar de un proceso abductivo o de una explicación abductiva. No obstante, ambas cosas están intrínsecamente relacionadas, de tal forma que una explicación abductiva es el producto de un proceso abductivo y un proceso abductivo da lugar a una explicación abductiva. La explicación o argumento abductivo puede definirse de la siguiente manera (Cfr. Aliseda; 2006, pp. 40-41): dada una teoría θ (conjunto de fórmulas) y una fórmula φ (fórmula atómica), α es una explicación abductiva si:

1. $\theta \cup \alpha \models \varphi$
2. α es consistente con θ
3. α es „mínima“

²¹ Una objeción muy común al hablar de comunidades científicas es apelar al relativismo y decir que como el conocimiento es relativo a cada comunidad científica no tiene sentido hablar de éste. Sin embargo, esto se evita si se acepta un pluralismo teórico. El pluralismo teórico permite que las comunidades científicas compartan bases teóricas comunes y que tenga sentido hablar de conocimiento entre las distintas comunidades científicas, pues se establecen los criterios de evaluación, los problemas relevantes y los valores epistémicos, entre otros, que se compartirán en un sistema y marco conceptual común. En cambio, el relativismo es cerrado y mantiene a los elementos implicados separados y ajenos.

4. α tiene una forma sintáctica restringida (usualmente una fórmula atómica o una conjunción de ellas).

Atendiendo a lo anterior, también es posible decir que los experimentos mentales pueden ser considerados como *razonamientos abductivos* y como *argumentos o estructuras abductivas*; es decir, que los experimentos mentales -experimentación mental- pueden ser considerados como *procesos abductivos*, donde el científico o el filósofo exploran diferentes escenarios antes de intentar llevarlos a cabo, ejecutando una serie de acciones imaginarias para generar posibles explicaciones, y que los experimentos mentales pueden ser considerados como *estructuras abductivas* o *argumentos de plausibilidad*, ya que una vez producida una explicación, esto es, una vez que se ha elegido una explicación como la más plausible de todo el conjunto de explicaciones posibles, ésta se vuelve una herramienta de argumentación. A través del experimento mental el científico argumenta, es decir, ofrece razones a favor o en contra de una determinada hipótesis o teoría en la ciencia. Los experimentos mentales como estructuras abductivas son únicamente plausibles, no concluyentes.

Para los fines de este trabajo me centraré en lo último, a saber, en la idea de que algunos experimentos mentales pueden ser reconstruidos como argumentos o estructuras abductivas.

4.2 Abducción e inferencia a la mejor explicación.

La inferencia a la mejor explicación es una forma de razonamiento no deductivo propuesta por Gilbert Harman (1965) y consiste en elegir la mejor explicación²² de un conjunto de explicaciones potenciales, dada la evidencia relevante; es decir aquella que explique de la mejor manera y de acuerdo a los criterios de evaluación aceptados un hecho o fenómeno. Su estructura puede ser reconstruida de la siguiente forma:

Dada la evidencia E

Y las explicaciones $A_1, [\dots] A_n$ de E

²² Para Harman la mejor explicación debe ser capaz de explicar la evidencia y al mismo tiempo superar a todas las demás hipótesis alternativas. Sin embargo, mismo Harman admite que determinar cuál hipótesis es "mejor" que otra es todo un problema, ya que las consideraciones y criterios de juicio varían considerablemente. Así, algunas consideraciones se enfocan en cuál hipótesis es la más simple, cuál es la más abarcadora, cuál es menos *ad hoc*, etc.

Es posible inferir la explicación del conjunto que mejor explique, en función de los criterios de evaluación aceptados por la comunidad científica, a E

$$\begin{array}{l} E \\ A_1 \rightarrow E \\ \dots \\ A_n \rightarrow E \\ / \text{ (Mejor) } A_i \end{array}$$

Al hacer este tipo de inferencia, y aceptando el supuesto que alguna de las hipótesis explicaría la evidencia, se infiere la mejor hipótesis del conjunto de hipótesis. Ésta debe ser capaz de excluir a todas las demás hipótesis alternativas y ser la que provea la “mejor” explicación dada la evidencia. Sin embargo, determinar cuál es la “mejor explicación”, ya sea que se trate de una hipótesis o teoría, no es tan simple.

Harman reconoce que no es fácil explicar la naturaleza de las distintas consideraciones que se toman en cuenta a la hora de evaluar el conjunto de hipótesis o explicaciones potenciales, ya que esto dependerá de los valores epistémicos o criterios de elección que cada sujeto o comunidad científica considere y priorice. Por lo general, las consideraciones se basan en cuál hipótesis es más simple, cuál explica más, cuál es menos *ad hoc*, o cuál es más plausible.

Para Thagard (1978) por ejemplo, los criterios que se deben considerar para la evaluación de teorías son tres:

1. **Comprehensividad:** se entiende como una medida para determinar qué tanto explica una teoría; una teoría será más comprehensiva o conciliadora que otra en la medida que explique más clases de hechos dada la evidencia relevante.
2. **Simplicidad:** una teoría será más simple en la medida que utilice menos hipótesis *ad hoc* o hipótesis auxiliares; éstas son proposiciones que no son parte de la teoría original pero que se asumen para ayudar a explicar un elemento “x” o un rango de hechos.
3. **Analogía:** es una herramienta de comparación, apoya a las teorías mejorando sus explicaciones; se señalan las características generales entre dos objetos y se generan razonamientos basados en las semejanzas, con la finalidad de explicar un objeto a través de las propiedades del otro.

La idea no es “aceptar” una hipótesis o teoría únicamente porque “hay algo que explica”, sino “aceptar” la hipótesis o teoría porque provee la “mejor explicación” dada la evidencia disponible.

[...] inference to the best explanation consists in accepting a hypothesis on the grounds that it provides a better explanation of the evidence than is provided by alternative hypothesis.

(Thagard, 1978; p. 77)

Para Thagard la mejor explicación se determina por estos tres criterios; cabe mencionar que éstos no fungen como condiciones necesarias y suficientes, sino como estándares de juicio considerados más sólidos que otros. Si bien estos criterios han sido propuestos para la evaluación de teorías es posible trasladarlos, al menos los dos primeros, a la evaluación de las explicaciones potenciales.

De esta manera, la inferencia a la mejor explicación y la abducción son dos procesos de razonamiento que tienen varias similitudes. Harman prácticamente los identifica como procesos similares:

“The inference to the best explanation” corresponds approximately to what others have called “abduction”, “the method of hypothesis”, “hypothetic inference”, “the method of elimination”, “eliminative induction”, and “theoretical inference”.

(Harman, 1965; pp. 88-89)

Sin embargo, así como se identifican ciertas similitudes, se identifican también algunas diferencias. Estas son algunas de ellas:

Algunas similitudes entre abducción e inferencia a la mejor explicación:

1. Ambas son formas de razonamiento ampliativo.
2. Se enfocan en la generación de explicaciones potenciales.
3. Proceden de manera retrospectiva.

Algunas diferencias:

1. Puede considerarse que la abducción se refiere más al “proceso de descubrimiento”, es decir, al proceso por el cual se obtienen explicaciones potenciales o diversas hipótesis explicativas dada la evidencia relevante.

2. Y que la inferencia a la mejor explicación se ocupa más de los “criterios de selección” que deben aplicarse para determinar cuál de las diversas hipótesis explicativas es la mejor.

De esta manera, es posible considerar que mientras la abducción se enfoca más en el proceso generativo de ideas o explicaciones, la inferencia a la mejor explicación se preocupa más por los criterios epistémicos de elección. En cualquier caso, considero que ambos tipos de razonamiento constituyen, en última instancia, terminologías diferentes para un mismo proceso metodológico que aprovecha sus similitudes y unifica sus diferencias. Sin embargo, no es mi intención abordar esta cuestión, que es por demás interesante, sino simplemente señalar que ambos tipos de razonamientos se complementan; así la “mejor explicación” en el razonamiento abductivo será, como ya se mencionó, la explicación considerada como la más plausible, es decir, aquella que sea más creíble, confiable y convincente para un sujeto o una comunidad científica en particular y en función de los valores de elección aceptados por esa comunidad y de la evidencia disponible.

5. Aplicación.

A continuación se realizan tres reconstrucciones de experimentos mentales desde la perspectiva abductivista. Los experimentos mentales a reconstruir son el experimento mental del reloj en la caja entre Einstein y Bohr, el experimento mental de Galileo y las balas de cañón y el experimento mental de J. S. Bell y las naves espaciales. En cada uno de los tres casos se derivan dos conclusiones distintas e incluso contrarias y resulta interesante señalar que mientras en el primer caso las explicaciones se producen a partir de dos teorías distintas, en el segundo y tercer caso las explicaciones se derivan a partir de la misma teoría pero de interpretaciones distintas de ésta. El objetivo de estas reconstrucciones es mostrar que algunos experimentos mentales pueden ser reconstruidos como estructuras abductivas o argumentos de plausibilidad, con todas las características que esto implica, y que constituyen herramientas metodológicas para evaluar, fortalecer o debilitar hipótesis o teorías en la ciencia.

5.1 El experimento mental del reloj en la caja entre Einstein y Bohr.

El experimento mental del reloj en la caja fue diseñado por Einstein para argumentar en contra de la teoría cuántica desarrollada por Heisenberg y Schrödinger, entre otros; su intención era mostrar que se podía medir con exactitud la velocidad y posición de una partícula al mismo tiempo y por lo tanto, que el principio de incertidumbre de Heisenberg era falso.

El principio de incertidumbre de Heisenberg afirma que no es posible conocer la velocidad y la posición de una particular subatómica al mismo tiempo, ya que existe un límite irreducible cuando se intenta medir con precisión la conjugación de un par de variables.

The uncertainty principle refers to the degree of indeterminateness in the possible present knowledge of the simultaneous values of various quantities with which the quantum theory deals; it does not restrict, for example, the exactness of a position measurement alone. Thus suppose that the velocity of a free electron is precisely known, while the position is completely unknown. Then the principle states that every subsequent observation of the position will alter the momentum by an unknown and undeterminable amount such that after carrying out the experiment our knowledge of the electron motion is restricted by the uncertainty relation. This may be expressed in concise and general terms by saying that every experiment destroys some of the knowledge of the system which was obtained by previous experiments.

(Heisenberg, 1930; p. 20)

El experimento mental consiste en imaginar una caja llena de fotones, ésta tiene en una de sus paredes un orificio que se controla por un dispositivo y cuando el orificio se abre escapa un solo fotón. Después se pide que se pese la caja para saber la diferencia del peso, comparando el peso inicial de la caja con todos los fotones adentro y el peso después de que escapa el fotón. Una vez que se obtiene la masa del fotón se utiliza la ecuación de Einstein $E=mc^2$ para calcular la energía del fotón. De esta manera, Einstein afirma que es posible calcular con exactitud la velocidad y saber la posición de una partícula al mismo tiempo y por tanto, que el principio de incertidumbre de Heisenberg es falso. (Cfr. Bishop, 1999; p. 536)

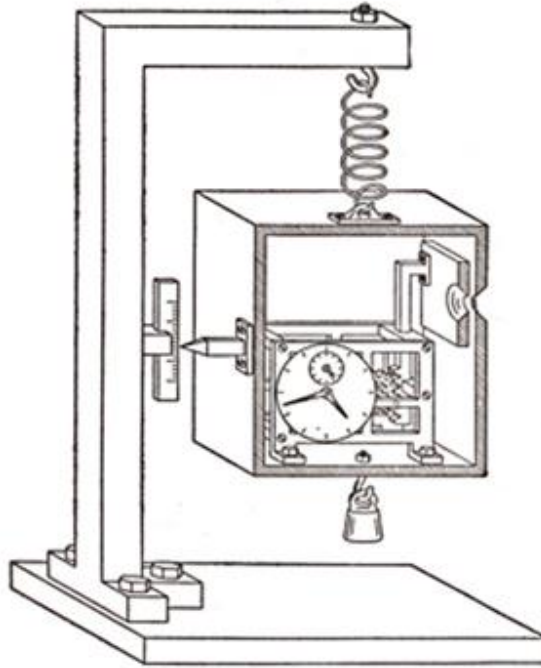


Fig. 1

Bohr usando el mismo experimento mental argumenta a favor del principio de incertidumbre de Heisenberg y muestra que no es posible medir con exactitud la velocidad y la posición de una partícula al mismo tiempo, sino únicamente una de las dos. Como se mencionó en el capítulo 3, Bohr se enfoca en los instrumentos prácticos y procedimientos que se tendrían que utilizar para medir la energía del fotón en un tiempo particular y muestra que “pesar” la caja requeriría que el aparato del reloj en la caja se mueva en un campo gravitacional, lo cual implica un límite fundamental en la precisión con el que el aparato del reloj en la caja puede medir el peso del fotón (y por lo tanto su energía) en un tiempo particular. Y de esta manera, reivindica el principio de incertidumbre de Heisenberg.

Así, del experimento mental del reloj en la caja se derivan dos conclusiones o explicaciones distintas y contrarias:

E₁) Es posible medir con exactitud la velocidad y la posición de una partícula al mismo tiempo.

E₂) No es posible medir con exactitud la velocidad y la posición de una partícula al mismo tiempo.

Einstein afirma E₁) y Bohr afirma E₂). El experimento mental del reloj en la caja sirve como un recurso metodológico para argumentar a favor o en contra de una teoría

determinada; en este caso, Einstein hace uso del experimento mental para argumentar en contra del principio de incertidumbre de Heisenberg y en contra de la teoría cuántica de aquella época y Bohr usa el experimento mental para argumentar a favor.

De acuerdo con la estructura abductiva el experimento mental puede ser reconstruido de la siguiente manera.

C = Fenómeno o hecho a explicar: velocidad y posición de una partícula subatómica. ¿Es posible medirlas al mismo tiempo?

E = Conjunto de explicaciones posibles y plausibles: E_1 es la explicación de Einstein y E_2 es la explicación de Bohr.

C

E_1

E_2

... $E_2 \rightarrow C$

\therefore Plausiblemente E_2

Einstein

[Teoría] Interpretación einsteniana de la mecánica cuántica.

[Fenómeno] La velocidad y posición de una partícula subatómica. ¿Es posible medir la velocidad y posición de una partícula al mismo tiempo?

[Explicación] Sí es posible medir con exactitud la velocidad y posición de una partícula al mismo tiempo. La indeterminación (Principio de incertidumbre de Heisenberg) indica que la función de onda no da una descripción completa de una partícula, ésta representa un conjunto de partículas y las predicciones probabilísticas derivadas de su posición; por tanto, es posible tener conocimiento preciso de la magnitud de las cantidades complementarias, lo cual significa una violación al principio de incertidumbre, es decir, su falsedad.

Bohr

[Teoría]	Interpretación de Copenhague de la mecánica cuántica (Heisenberg y Bohr entre otros)
[Fenómeno]	La velocidad y posición de una partícula subatómica. ¿Es posible medir la velocidad y posición de una partícula al mismo tiempo?
[Explicación]	No es posible medir con exactitud la velocidad y posición de una partícula al mismo tiempo. Una partícula puede manifestar ambas cosas, tanto una posición definida o un momento definido pero no es posible medirlas al mismo tiempo. La función de onda sí da una descripción completa del estado de una partícula.

El experimento mental del reloj en la caja puede ser concebido como un argumento o estructura abductiva porque da lugar a dos conclusiones o explicaciones distintas y en este caso contrarias, que se derivaron a partir de supuestos o teorías distintas. Y porque estas conclusiones o explicaciones no son concluyentes, sino únicamente plausibles, la aparición de nueva información, desarrollo de nuevas tecnologías, reinterpretación de datos, modificación de la teoría, etc., puede invalidar las conclusiones derivadas hasta el momento y generar nuevas explicaciones; aunque como ya se mencionó esto aplica para todo tipo de explicación. De esta forma, el experimento mental al ser considerado un argumento o estructura abductiva no está obligado a derivar con necesidad una única conclusión como los argumentos deductivos, sino una serie de posibles y plausibles conclusiones o explicaciones; en este caso, del experimento mental del reloj en la caja se derivaron dos conclusiones o explicaciones distintas: E_1 y E_2 .

Tanto Einstein y Bohr están intentando argumentar a través del mismo experimento mental, uno en contra del principio de incertidumbre de Heisenberg y de la teoría cuántica y otro a favor; ambos están intentando ofrecer razones para fortalecer sus hipótesis respectivamente.

Así, es posible observar que el experimento mental del reloj en la caja constituye:

- Una herramienta metodológica utilizada para fortalecer una hipótesis, en este caso, Einstein quería fortalecer la hipótesis E_1) y Bohr la hipótesis E_2). O atacar o fortalecer una teoría en particular, la teoría que está de fondo es la teoría cuántica de aquella época, Einstein quería atacarla y Bohr defenderla.

- Una estructura abductiva, ya que se derivan dos conclusiones distintas y en este caso contrarias, las cuales no son concluyentes, sino únicamente plausibles. La explicación de Bohr es considerada más plausible que la de Einstein porque no genera inconsistencias con la teoría desde la cual se genera la explicación para el fenómeno planteado, pero el desarrollo de la ciencia puede invalidar dicha explicación y mostrar a través de nuevas teorías mejores explicaciones.

5.2 El experimento mental de Galileo y las balas de cañón.

Galileo en el *Discorsi* argumentó que todos los cuerpos independientemente de su peso caen a la misma velocidad (Galileo, 1974). Galileo comienza por hacer notar que para Aristóteles los cuerpos pesados caen más rápido que los cuerpos livianos o menos pesados ($P > L$). Luego nos pide que imaginemos que una bala de cañón pesada se une a una bala de mosquete ligero y se pregunta ¿qué pasaría si se liberarán juntas?



Fig. 2

Desde la teoría aristotélica del movimiento la bala de mosquete ligera hará más lenta a la bala de cañón pesada, actuando como una especie de arrastre, por lo que la velocidad del sistema combinado ($P+L$) sería más lenta que la velocidad de la bala de cañón pesada que cae sola ($P > P+L$). Por otro lado, el sistema combinado es más pesado que la bala de cañón pesada, por lo que debe caer más rápido ($P+L > P$). De esta manera, desde la teoría aristotélica del movimiento se llega a una contradicción: la bala de cañón pesada es más rápida y menos rápida que el sistema combinado. Así, la teoría aristotélica colapsa.

Entonces ¿cuál cae más rápido? Para Galileo la contradicción se resuelve haciendo que sean iguales, afirmando que todos los cuerpos caen a la misma velocidad (en el vacío) ($P=L=P+L$).

A partir del experimento mental de las balas de cañón se generan dos conclusiones o explicaciones distintas:

E₁) La bala de cañón más pesada cae más rápido y menos rápido que el sistema combinado. (Contradicción derivada desde la teoría aristotélica.)

E₂) La bala de cañón más pesada cae a la misma velocidad que la bala de mosquete más ligera y que el sistema combinado. ($P=L=P+L$).

El experimento mental de Galileo acerca de la caída de los cuerpos muestra que la teoría aristotélica del movimiento es lógicamente imposible ya que deriva una contradicción, a saber, que la bala de cañón pesada es más rápida y menos rápida que el sistema combinado. De esta manera, Galileo muestra que la teoría de Aristóteles sobre la caída de los cuerpos falla.

Reconstrucción:

C = Fenómeno o hecho a explicar: caída de los cuerpos.

E = Conjunto de explicaciones posibles y plausibles: E₁ es la explicación de Aristóteles y E₂ es la explicación de Galileo.

C

E₁

E₂

... E₂ → C

∴ Plausiblemente E₂

Aristóteles

[Teoría] Teoría aristotélica del movimiento. Los cuerpos pesados caen más rápido que los cuerpos ligeros (o menos pesados).

[Fenómeno] Caída de los cuerpos. Si se libera una bala de cañón pesada atada a una bala de mosquete ligero ¿Qué pasaría?

[Explicación] La bala de cañón pesada (P) caerá más rápido que la bala de mosquete ligero (L). Al liberarse juntas (P+L) se tiene que la bala de mosquete ligero hará más lenta a la bala de cañón pesada, por lo que la velocidad del sistema combinado (P+L) sería más lenta que la velocidad de la bala de cañón pesada que cae sola ($P > P+L$). Por otro lado, el sistema combinado (P+L) es más pesado que la bala de cañón pesada sola, por lo que debe caer más rápido ($P+L > P$). Así se tiene que ($P > P+L$) y que ($P+L > P$).

Galileo

[Teoría] Teoría galileana del movimiento. Todos los cuerpos caen a la misma velocidad (en el vacío).

[Fenómeno] Caída de los cuerpos. Si se libera una bala de cañón pesada atada a una bala de mosquete ligero ¿Qué pasaría?

[Explicación] El sistema combinado (P+L) cae a la misma velocidad que las balas de cañón por separado. ($P=L=P+L$).

En la estructura:

C

E_1

E_2

... $E_2 \rightarrow C$

∴ Plausiblemente E_2

E_2 es considerada más plausible que E_1 porque no lleva a una contradicción. E_1 y E_2 dependen de la teoría desde la cual se interpreta el fenómeno u observación, en este caso E_1 se genera desde la teoría aristotélica y E_2 desde la teoría galileana del movimiento. Sin embargo, en ocasiones la teoría desde la cual se generan las posibles explicaciones es la misma, como en el experimento mental de J. S. Bell sobre relatividad especial que veremos en la siguiente sección.

La explicación de Galileo es más convincente que la explicación de Aristóteles acerca de la caída de los cuerpos; su explicación para el fenómeno de la caída de los cuerpos es más plausible que la teoría de Aristóteles, pues las razones que ofrece son más convincentes ya que no generan una contradicción como en Aristóteles.

El experimento mental de las balas de cañón puede ser considerado un argumento o estructura abductiva ya que se pueden derivar dos conclusiones o explicaciones distintas. Si bien no se trata de una discusión propiamente entre Aristóteles y Galileo, como en el caso del experimento mental del reloj en la caja entre Einstein y Bohr, es posible reconstruir igualmente el experimento mental.

Las dos conclusiones o explicaciones son:

E₁) La bala de cañón más pesada cae más rápido y menos rápido que el sistema combinado. (Contradicción derivada desde la teoría aristotélica).

E₂) La bala de cañón más pesada cae a la misma velocidad que la bala de mosquete más ligera y que el sistema combinado. ($P=L=P+L$)

La explicación considerada más plausible o convincente es E₂) ya que la primera deriva una contradicción lógica.

Así, es posible observar que el experimento mental de las balas de cañón constituye:

- Una herramienta metodológica utilizada para fortalecer una hipótesis o teoría, en este caso, Galileo quería fortalecer la hipótesis E₂) y debilitar la hipótesis E₁). La teoría que Galileo quería fortalecer era su teoría del movimiento y debilitar o atacar la teoría aristotélica.
- Una estructura abductiva, ya que se derivan dos conclusiones distintas, las cuales no son concluyentes, sino únicamente plausibles. La explicación de Galileo es más plausible que la de Aristóteles porque no genera inconsistencias y porque el desarrollo de nuevas teorías en la ciencia puede invalidar dicha explicación.

5.3 El experimento mental de Bell y las naves espaciales.

En el libro *Speakable and Unspeakable*, cap. 9, "How to teach especial relativity" J. S. Bell propone un experimento mental sobre relatividad especial. Para Bell destacar la continuidad de las ideas tempranas sobre las nociones de espacio y tiempo con las ideas más actuales resulta de suma importancia cuando se enseña la relatividad especial, ya que la discontinuidad a menudo destruye por completo la confianza del estudiante en conceptos ya adquiridos. De esta manera, propone que imaginemos una situación que confronta a los estudiantes con la idea anterior:

Tres naves espaciales pequeñas, A, B y C se desplazan libremente por una región del espacio alejada de la influencia de cualquier otra materia, sin rotación y sin movimiento relativo, con B y C equidistante de A. (Fig. 3)



Fig. 3

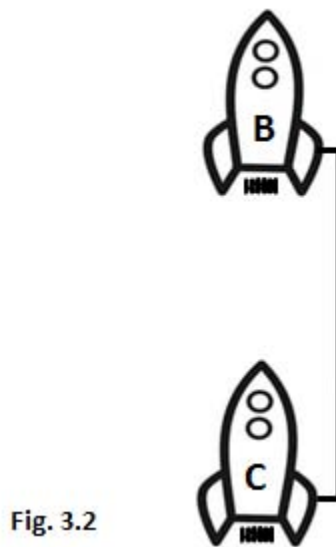
A la recepción de una señal de A los motores de B y C se encienden y aceleran suavemente. (Fig. 3.1)



Fig. 3.1

Supongamos que las naves B y C son idénticas y que tienen programas de aceleración idénticos. Entonces (según lo contado por un observador en A) tendrán en todo momento la misma velocidad y así seguirán desplazándose uno del otro por una distancia fija.

Supongamos ahora que una frágil cuerda (o hilo) se ata inicialmente entre las proyecciones de B y C. (Fig. 3.2)



Si la cuerda es lo suficientemente larga para cubrir la distancia requerida inicialmente, entonces a medida que los cohetes aceleren, se convertirá en demasiado corta, debido a su necesidad de contracción de Fitzgerald. La contracción de Fitzgerald (1889) es un efecto relativista que consiste en la contracción de la longitud de un cuerpo en la dirección de su movimiento a medida que su velocidad se acerca a la velocidad de la luz. De esta manera, la cuerda deberá romperse cuando, a una velocidad suficientemente alta, la contracción imponga estrés intolerable.

Bell se pregunta si esto ¿es realmente así? Y cuenta que este viejo problema vino a discusión una vez entre un físico experimental distinguido que se negó a aceptar que la cuerda se rompería. El físico experimental consideraba que la afirmación de Bell, de que la cuerda se rompería, se basaba en una interpretación incorrecta de la relatividad especial. Por lo que decidieron apelar a la División de arbitraje de Teorías del CERN e hicieron –afirma Bell- una consulta informal de la opinión. El consenso que surgió fue que la cuerda no se rompería.

Para Bell, muchas personas que dan esta respuesta errónea al principio, después de una reflexión más profunda obtienen la respuesta correcta. Así, las personas se

sienten obligadas, eventualmente, a trabajar cómo se ven las cosas para los observadores en B o C. Y encuentran que en B, por ejemplo, se ve a C a la deriva, y cada vez más atrás, de modo que una determinada pieza de cuerda no podría cubrir la distancia. Es sólo después de trabajar esto y quizás con un sentimiento de inquietud – afirma Bell- que esas personas finalmente aceptan que es perfectamente trivial la manera en cómo A ve las cosas. Para Bell los que tienen una formación más clásica y conocen algo del razonamiento de Larmor, Lorentz y Poincaré, así como de Einstein, tienen instintos más fuertes y más sólidos.

Así, a partir de este experimento mental se generan, igualmente, dos conclusiones distintas y contrarias:

E₁) La cuerda se rompería.

E₂) La cuerda no se rompería.

Bell afirma E₁) y el físico experimental y la División de arbitraje de Teorías del CERN afirma E₂).

El experimento mental de las naves espaciales de Bell es, igualmente, un recurso metodológico que sirve para evaluar, fortalecer o debilitar una hipótesis o teoría en la ciencia. En este caso, la teoría involucrada es la teoría de la relatividad especial y resulta interesante el hecho de que en este experimento mental no se esté argumentando a favor y en contra de esta teoría, sino de una interpretación en particular.

De acuerdo con la estructura abductiva este experimento mental puede ser reconstruido de la siguiente manera.

C = Fenómeno o hecho a explicar: El movimiento relativo de dos cuerpos, unidos por una cuerda, que viajan a una velocidad cercana a la de la luz. ¿La cuerda se rompería?

E = Conjunto de explicaciones posibles y plausibles: E₁ es la explicación de Bell y E₂ es la explicación del físico experimental y el consenso de la División de arbitraje de Teorías del CERN.

C

E₁

E₂

... $E_1 \rightarrow C$

\therefore Plausiblemente E_1

Cabe señalar que la explicación considerada como la más plausible (E_1 o E_2) dependerá de la interpretación que se adopte de la teoría en cuestión, ya que la situación planteada no deja claro quién ve la contracción, es decir, qué sistema inercial (A, B o C) se debe tener en cuenta.

Bell

[Teoría]	Teoría de la relatividad especial.
[Fenómeno]	El movimiento relativo de dos cuerpos, unidos por una cuerda, que viajan a una velocidad cercana a la de la luz. ¿La cuerda se rompería?
[Explicación]	<p>Si. Su conclusión se basa en la suposición de que los cuerpos físicos se contraen de manera relativa. Si la longitud de la cuerda es igual a la distancia inicial entre B y C, entonces a medida que los cohetes aceleren, se convertirá en demasiado corta, debido a su necesidad de contracción de Fitzgerald, y finalmente deberá romperse, ya que a una velocidad suficientemente alta la contracción impone estrés intolerable.</p> <p>La cuerda estará sometida a la contracción y, por tanto, no será capaz de cubrir la distancia necesaria con el tiempo y el aumento de velocidad.</p>

Físico experimental y la División de arbitraje de Teorías del CERN.

[Teoría]	Teoría de la relatividad especial. (Interpretación diferente)
[Fenómeno]	El movimiento relativo de dos cuerpos, unidos por una cuerda, que viajan a una velocidad cercana a la de la luz. ¿La cuerda se rompería?
[Explicación]	No. Desde la perspectiva de A la cuerda no se rompería. La distancia entre B y C se mantiene constante porque sus aceleraciones son idénticas. Como los motores son idénticos las naves no deberían atrasarse ni adelantarse una respecto de otra, sino avanzar como un sistema único.

El experimento mental de las naves espaciales de Bell puede ser concebido como:

- Una herramienta metodológica utilizada para fortalecer o debilitar una hipótesis. En este caso, Bell intenta fortalecer la hipótesis E_1) y el físico experimental y la División de arbitraje de Teorías del CERN la hipótesis E_2). La teoría que se intenta fortalecer es la teoría de la relatividad especial aunque desde dos interpretaciones distintas.
- Una estructura abductiva porque se derivan dos conclusiones o explicaciones distintas y porque éstas no son concluyentes, sino plausibles. La aparición de nueva información, reinterpretación de datos o incluso de la teoría misma, pueden invalidar las conclusiones derivadas hasta el momento y generar nuevas explicaciones.

6. Conclusiones.

Hemos defendido que los experimentos mentales constituyen herramientas metodológicas que sirven para evaluar, fortalecer o debilitar hipótesis, que a través de ellos el científico o el filósofo argumenta, es decir, ofrece razones a favor o en contra de una determinada teoría y que éstos pueden ser reconstruidos como argumentos o estructuras abductivas, esto es, como argumentos de plausibilidad, convincentes pero no concluyentes.

Para ello, revisamos de manera general las características que las diferentes propuestas sobre la naturaleza de los experimentos mentales consideran esenciales, como la cualidad *a priori*, el elemento modal, el carácter contrafáctico y la función heurística, entre otras, y consideramos que la función metodológica de los experimentos mentales constituye su característica más representativa y elemental, ya que antes que nada, los experimentos mentales constituyen herramientas metodológicas para la investigación.

Consideramos que los experimentos mentales como herramientas metodológicas constituyen una herramienta útil al servicio de la investigación científica y filosófica. Y que la discusión entre aquellos que están a favor o en contra de la utilidad de los experimentos mentales se puede explicar desde la modalidad, mostrando que cuando se prioriza la *posibilidad física* se tiende a rechazar la utilidad de los experimentos

mentales y en cambio, cuando se apela o prioriza la *posibilidad metafísica* se tiende a aceptar dicha utilidad.

También hemos argumentamos que los experimentos mentales pueden ser reconstruidos como argumentos o estructuras argumentativas. Vimos que algunos como Norton consideran que los experimentos mentales pueden ser reconstruidos, principalmente, como estructuras de tipo deductivo, aunque no descarta otros tipos de reconstrucción. Y que algunos como Bishop consideran que este tipo de reconstrucción colapsa, ya que en ciertos casos los experimentos mentales derivan más de una conclusión. En este punto hemos defendido que las reconstrucciones abductivas superan este problema y que pueden dar cuenta perfectamente de este tipo de casos, donde se deriva más de una conclusión.

Mostramos que el razonamiento abductivo es un tipo de razonamiento distinto al razonamiento deductivo e inductivo. Y que éste nos permite explicar un fenómeno o hecho observado, generando un conjunto de explicaciones posibles y eligiendo aquella explicación que el sujeto o la comunidad científica considere la explicación más plausible.

Finalmente las tres reconstrucciones abductivas de experimentos mentales que realizamos dan cuenta de casos donde la reconstrucción deductiva falla, porque se derivan dos conclusiones o explicaciones distintas e incluso contrarias. Las reconstrucciones abductivas muestran que los experimentos mentales constituyen herramientas metodológicas que el científico o el filósofo utilizan para evaluar, fortalecer o debilitar una hipótesis o teoría en la ciencia y que las conclusiones o explicaciones generadas no son definitivas, ya que la aparición de nueva información, reinterpretación de datos incluso de la teoría misma, pueden invalidar las conclusiones derivadas hasta el momento y generar nuevas y mejores explicaciones; en las reconstrucciones abductivas las conclusiones o explicaciones sólo se afirman en términos de plausibilidad. Así mismo, es posible afirmar desde la perspectiva abductivista que los experimentos mentales no sólo constituyen argumentos de plausibilidad, sino una manera legítima de argumentar en la ciencia.

Cabe mencionar que si bien este trabajo se enfocó principalmente en tres experimentos mentales de la física, los resultados podrían proyectarse a otras áreas del conocimiento, como la filosofía de la mente, la ética, las matemáticas, entre otras. Así mismo, las líneas de investigación que se desprenden de este trabajo son varias, destaco dos en particular: la cuestión de la relevancia científica y el impacto que los experimentos mentales han tenido en las prácticas de investigación. Ya que si

aceptamos que los experimentos mentales constituyen herramientas metodológicas para la investigación, cabe preguntarnos por las condiciones que un experimento mental necesita cumplir para que sea considerado científicamente relevante, porque es un hecho que existen muchos experimentos mentales en la ciencia, pero ¿son todos científica o filosóficamente relevantes? ¿En qué situaciones y bajo qué circunstancias resultan significativos los experimentos mentales? Y en relación con esto, preguntarnos ¿cuál ha sido su impacto en las prácticas de investigación, así como en las formas de razonamiento? Esto es, ¿cuáles han sido las implicaciones y consecuencias que a partir de su consolidación como herramientas metodológicas se han derivado en la investigación científica? Estas dos cuestiones están en estrecha relación con la función metodológica que hemos defendido aquí, por lo que resultaría de gran interés profundizar en ellas, no sólo para profundizar en el análisis, sino para enriquecer nuestras reflexiones sobre la naturaleza y el papel que los experimentos mentales desempeñan en la ciencia.

7. Referencias.

Adler, Jonathan E. and Rips, Lance J. (eds.) (2008), *Reasoning: Studies of Human Inference and its Foundations*, Cambridge University Press.

Aliseda Llera, Atocha (2006), *Abductive Reasoning. Logical Investigations into Discovery and Explanation*, (eds.) Vincent F. Hendricks, John Symons, Jaakko Hintikka [et. al.], Synthese Library: Studies in Epistemology, Logic, Methodology, and Philosophy of Science, vol. 330, Springer.

Bealer, George (1998), "Intuition and the autonomy of philosophy," in DePaul and W. Ramsey (eds.) *Rethinking Intuition: The Psychology of Intuition and Its Role In Philosophical Inquiry*, pp. 201-239.

Bell, J. S. (1987), "How to teach special relativity", in *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, pp. 67-80.

Bishop, Michael, (1999), "Why thought experiments are not arguments", *Philosophy of Sciences*, 66, pp. 534-541.

Bohr, Niels (1949), "Discussion with Einstein on Epistemological Problems in Atomic Physics", originally published in *Albert Einstein: Philosopher-Scientist*, (eds.) P. A. Schilpp, pp. 201-41, *The Library of Living Philosophers*, Evanston (1949).

Brown, James Robert (1991), *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. London: Routledge.

Duhem, Pierre (1914), *La théorie physique son objet-sa structure*, segunda parte, cap. vi, *La théorie physique et l'expérience*, pp. 146-176.

----- (1914), *The Aim and Structure of Physical Theory*. (Trans.) P. Wiener. Princeton: Princeton University Press, 1954.

Engel, Pascal (2011), "Philosophical Thought Experiments: In or Out of the Armchair?", in Roux and Ierodiakonou, pp. 145-164.

Forge, John (1991), "Thought Experiments in the Philosophy of Physical Science", in Horowitz and Massey, pp. 209-222

Gigerenzer, Gerd (2004), "Fast and Frugal Heuristics: The tools of Bounded Rationality", en *Blackwell Handbook of Judgment and Decision Making*, (eds.) Derek J. Koehler and Nigel Harvey, Ed. Blackwell Publishing, pp. 62-88.

Gooding, David D. (1992), "What is Experimental about Thought Experiments?", *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2, (eds). David Hull, Micky Forbes and Kathleen Okruhlik, pp. 280-290.

Hacking, Ian (1992), "Do Thought Experiments Have a Life of Their Own? Comments on James Brown, Nancy Nersessian and David Gooding", *Proceeding of the Biennial Meeting of the Philosophy of Science Association*, vol. 2, pp. 302-308.

Hanson, N.R. (1972), *Patterns of Scientific Discovery*. Cambridge at The University Press.

Harman, Gilbert (1965), "The Inference to the Best Explanation", *The Philosophical Review*, Published by Duke University Press on behalf of Philosophical Review, vol. 74, No. 1, pp. 88-95.

Heisenberg, W. (1930), *The Physical Principles of the Quantum Theory*, Chicago: University of Chicago Press.

Horowitz, Tamara and Gerald J. Massey (eds.) (1991), *Thought Experiments in Science and Philosophy*. Savage, MD: Rowman & Littlefield Publishers.

Irvine, Andrew D. (1991), "On the Nature of Thought Experiments in Scientific Reasoning", in Horowitz and Massey, pp. 149-166.

Kant, Immanuel (2002), *Crítica de la razón pura*, tr. Manuel García Morente, Madrid, ed. Tecnos.

Koyré, Alexandre, (1968), *Metaphysics and Measurement*, London: Chapman & Hall.

Kuhn, Thomas (1982), *La tensión esencial. Estudios selectos sobre la tradición y el cambio en el ámbito de la ciencia*, tr. Roberto Helier, México, FCE.

Lakatos, Irme (1989), "La historia de la ciencia y sus reconstrucciones racionales", en *La metodología de los programas de investigación científica*, ed. Alianza Universidad, editado por John Worall y Gregory Currie, versión española de Juan Carlos Zapatero y revisión de Pilar Castrillo, Madrid, pp. 134-179.

Mach, Ernst (1948), *Conocimiento y Error*, tr. Cortés Pla. Espasa-Calpe, Buenos Aires, Argentina, S. A.

McAllister, James, W. (2012), "Thought experiments and the exercise of Imagination in Science", en Brown, J. et. al. (eds.), *Thought Experiments in Science, Philosophy and the Arts*, pp. 11-29.

Nersessian, Nancy (1992), "In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling", *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1992, vol. 2, pp. 291-301.

Norton, John D. (1991), "Thought experiments in Einstein's work", in Horowitz and Massey, pp. 129-148.

----- (2004), "Why Thought Experiments Do Not Transcend Empiricism", in Christopher Hitchcock (ed.), *Contemporary Debates in the Philosophy of Science*. Oxford: Blackwell, pp. 44-66.

Peirce, C. S. (1958), *Collected Papers of Charles Sanders Peirce*. Volumes 1–6 edited by C. Hartshorne, P. Weiss, Cambridge, Harvard University Press, 1931–1935; and volumes 7–8 edited by A.W. Burks. Cambridge, Harvard University Press. 1958.

Pollock, John L. (2008), "Defeasible Reasoning", in Adler and Lance, pp. 451-470.

Rappaport, Steven (1996), "Inference to the Best Explanation: Is It Really Different from Mill's Methods?", *Philosophy of Science*, Published by The University of Chicago Press on behalf of the Philosophy of Science, vol. 63, No. 1, pp. 65-80.

Reichenbach, Hans. (1938), *Experience and Prediction*, Chicago: University of Chicago Press.

Rescher, Nicholas (1991), "Thought Experimentation in Presocratic Philosophy", in Horowitz and Massey, pp. 31-42.

Roux, Sophie y Ierodiakonou, Katerina (eds.) (2011), *Thought Experiments in Methodological and Historical Contexts*. Brill, Leiden, Boston.

Sorensen, Roy (1992), *Thought Experiments*, Oxford, Oxford University Press.

Thagard, Paul (1978), "The Best Explanation: Criteria for Theory Choice", *The Journal of Philosophy*, Published by Journal of Philosophy, Inc., vol. 75, No. 2, pp. 76-92.

----- (2012), Thought Experiments considered harmful, University of Waterloo,
Draft 4, May, 2013.

Williamson, Timothy (2007), *The Philosophy of Philosophy*, Oxford: Blackwell.