

ISSN 2525-1198

Volumen 1
Nº 1
Octubre 2016

Epistemología e Historia de la Ciencia

Área Lógico-Epistemológica de la Escuela de Filosofía,
Centro de Investigaciones de la Facultad de Filosofía y Humanidades,
Universidad Nacional de Córdoba



Epistemología e Historia de la Ciencia

Epistemología e Historia de la Ciencia es una revista digital, de aparición semestral, dedicada a la publicación de artículos originales de filosofía general de la ciencia y filosofías de las ciencias particulares, así como artículos de historia de la ciencia con orientación filosófica. Las áreas de interés son entendidas en un sentido amplio y teóricamente plural.

Editorial y correspondencia

Revista Epistemología e Historia de la Ciencia
Centro de Investigaciones Facultad de Filosofía y Humanidades (CIFFyH), Pabellón Agustín Tosco, Ciudad Universitaria, Córdoba (5000), Argentina. Universidad Nacional de Córdoba.

Información adicional y envío de artículos

<https://revistas.unc.edu.ar/index.php/afjor/index>

Correo electrónico

revistaepistemologia@ffyh.unc.edu.ar

Comité editorial

Editora Responsable

Marisa Velasco, Universidad Nacional de Córdoba (Argentina)

Editores

Pío García, Universidad Nacional de Córdoba (Argentina)

Luis Salvático, Universidad Nacional de Córdoba (Argentina)

A. Nicolás Venturelli, Universidad Nacional de Córdoba, CONICET (Argentina)

Hernán Severgnini, Universidad Nacional de Córdoba (Argentina)

Comité Académico

Mario Casanueva, Universidad Autónoma Metropolitana (México)

Silvio Seno Chibeni, Departamento de Filosofía, Unicamp (Brasil)

Miguel Angel Fuentes, Instituto de Sistemas Complejos (Chile), Santa Fe Institute (Estados Unidos)

Lucía Lewowicz, Universidad de la Republica (Uruguay), Max Planck Institute for the History of Science (Alemania)

Oswaldo Pessoa Jr., Departamento de Filosofía, Universidade de São Paulo (Brasil)

Anna Carolina K.P. Regner, Instituto Latino Americano de Estudos Avançados Universidade Federal do Rio Grande do Sul (Brasil)

Víctor Rodríguez, Universidad Nacional de Córdoba (Argentina)

Secretario

Julián Reynoso

Organismo Responsable

Área de Filosofía del Centro de Investigaciones de la Facultad de Filosofía y Humanidades y Escuela de Filosofía de la Universidad Nacional de Córdoba.

Pabellón Agustín Tosco, Ciudad Universitaria, Córdoba Capital.

ISSN: 2525-1198

Índice

Editorial	5
Hernán Miguel	
Similaridad entre mundos posibles: un viaje de ida	7
Hernán Lucas Accorinti, Juan Camilo Martínez González	
Modelos y representación	21
Marcela Becerra Batán	
La cuestión de la Epistemología Histórica como estilo epistemológico	35
Diego Lawler	
Actuar intencionalmente: esbozo de una teoría de la intención	53
Itatí Branca	
Reseña: <i>Simulation and Similarity</i> de Michael Weisberg.....	87

Editorial

Con gran satisfacción presentamos el primer número de la revista *Epistemología e Historia de la Ciencia*. Esta publicación ha significado un esfuerzo importante que intenta ser una continuación de la tradición inaugurada en Córdoba (Argentina) por las Jornadas de Epistemología e Historia de la Ciencia, organizadas por el Área Lógico-Epistemológica de la Escuela de Filosofía y el Centro de Investigaciones María Saleme de Burnichon de la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba (Argentina). Hace 27 años, por iniciativa del profesor Víctor Rodríguez, estas Jornadas comenzaron a realizarse, contando con el apoyo de referentes del área, muy especialmente del profesor Gregorio Klimovsky. Nacieron como una pequeña reunión de carácter local para paulatinamente constituirse en un espacio de discusión e intercambio de ideas en las áreas de filosofía e historia de la ciencia del Cono Sur.

En los últimos 15 años se presentaron más de cien trabajos en cada edición, incluyendo mesas sobre temáticas específicas, en torno a un abanico amplio de temas vinculados a la epistemología y la historia de la ciencia y de disciplinas científicas particulares. En los inicios de las Jornadas se publicaba un volumen de actas de los trabajos expuestos. Con el tiempo comenzó a editarse un volumen con una selección de trabajos con editores invitados y con una evaluación ciega por pares. Se publicaron un total de 22 volúmenes, incluidas las dos primeras actas. En 2015 se publicó el último volumen con este formato en una edición conjunta con la Asociación de Filosofía e Historia de la Ciencia del Cono Sur (AFHIC).

Por la modalidad y estilo propio de las Jornadas, esta reunión sigue siendo un punto de encuentro que no sólo se destaca por sus aspectos académicos sino también por su dimensión humana. En todos estos años una comunidad de colegas ha contribuido de manera significativa, aportando sus saberes y su apoyo. Entre ellos queremos recordar especialmente a Horacio Faas, Eduardo Flichman, Guillermo Boido, Luly Horenstein, Alberto Moreno, Gladys Palau y Pablo García.

Es nuestro deseo que el espíritu de la tradición inaugurada hace casi tres décadas, y aún vigente en nuestras reuniones anuales, se recree y profundice en la publicación que presentamos aquí. Con ella, queremos contribuir a consolidar el debate de problemáticas epistemológicas en las comunidades de Latinoamérica. A este objetivo responde la decisión de solo aceptar trabajos en español y en portugués.

Epistemología e Historia de la Ciencia es una revista digital, de aparición semestral, dedicada a la publicación de artículos originales de filosofía general de la ciencia y filosofías de las ciencias particulares, así como artículos de historia de la ciencia con orientación filosófica. Las áreas de interés son entendidas en un sentido amplio y teóricamente plural.

Queremos agradecer muy especialmente a los evaluadores externos y a los

autores que contribuyeron a este primer número por confiar en este proyecto. Invitamos cordialmente a la comunidad académica a colaborar con la difusión de esta propuesta y a participar en los futuros números.

Marisa Velasco
Pío García
Nicolás Venturelli
Luis Salvatico
Hernán Severgnini
Editores - *Epistemología e Historia de la Ciencia*

Similaridad entre mundos posibles: un viaje de ida

Hernán Miguel¹

Recibido: 15 de abril de 2016.

Aceptado: 26 de julio de 2016.

Resumen. Se suele aceptar que no es posible garantizar la simetría de la relación de similaridad comparativa global entre mundos posibles. Sin embargo, existen usos exitosos y tentadores al suponer casos simétricos que pueden dar lugar a pensar que la falta de garantía se debe a cuestiones formales. Es decir, a pesar de que la relación no es simétrica, se la usa como simétrica ya que la gran mayoría de los ejemplos se ajustan a esta simetría “caso”. En este trabajo presento un tipo de falla en la simetría de la similaridad entre mundos posibles y expongo las características por las cuales se produce esa falla. Esta caracterización es útil también para dar cuenta del porqué del aparente éxito en suponer simétrica esa relación para los casos que no presentan estas características. A continuación muestro las consecuencias que tiene esta falla en los usos que hace Phil Dowe de la relación de causación para dar cuenta de los casos que involucran omisiones, casos en que se hace indispensable para su teoría el uso de contrafácticos.

Palabras clave: similaridad – mundos posibles – causación – omisiones.

Title: Similarity between possible worlds: A one way trip.

Abstract. It is widely admitted that is not possible to warrant the symmetry for the global comparison similarity relation between possible worlds. Despite this, there are so many successfully analyzed cases treated as symmetrical so it rise the temptation to regard the problem as a formal one, in the sense of the conditions are not fulfilled for the symmetry be taken for granted. In this work I show a kind of failure in the symmetry of the similarity relation between possible worlds and state the features that are responsible for that failure. This approach become useful also for giving account of the apparently unproblematic cases where the symmetry seems to apply smoothly. Later, I analyze how the Phil Dowe’s theory of causation is affected by this kind of failures at the time to face the cases of causation with omission, where the theory cannot avoid the counterfactual analysis, as developed by the own author.

Keywords: similarity – possible worlds – causation – omissions.

1. Introducción

David Lewis (1973) en su *Counterfactuals* ofrece como criterio de evaluación para un condicional contrafáctico el análisis de lo que ocurre en los mundos posibles más similares al mundo efectivo, en los que se cumple el antecedente del condicional contrafáctico. Dicho brevemente, para saber si es verdadero o falso el contrafáctico “si no hubiera ocurrido A, no habría ocurrido B” debemos buscar de entre todos los mundos posibles en los que no ha ocurrido A, aquellos más similares a nuestro mundo, en el que

¹ Universidad de Buenos Aires / SADAF.

✉ ciencias@retina.ar

Miguel, Hernán (2016). Similaridad entre mundos posibles: un viaje de ida. *Epistemología e Historia de la Ciencia*; 1(1); 7-20. ISSN: 2525-1198



A efectivamente ocurrió. Al buscar en ellos debemos inspeccionar si en los más similares a nuestro mundo ocurre B o no ocurre. Si en todos los mundos más similares a partir de cierto umbral, ocurre B, entonces el condicional es verdadero. Si entre de ellos, de los más similares, siempre se puede encontrar alguno en el que B no ocurre, entonces el condicional es falso.

Tal criterio es establecido por Lewis con gran cuidado de evitar hacer mención de “el mundo más similar” ya que bien podría ocurrir que el conjunto de los mundos más similares no tuvieran un máximo, es decir, podría no existir uno y solo uno identificable con el más cercano. Podría darse el problema que el mismo Lewis advierte como el supuesto de límite.²

Este cuidado de referirse a los más similares y no al más similar fue un punto de partida sobre el cual edificar la advertencia de que, si cierto mundo parecía ser el más similar al nuestro, no necesariamente el nuestro sería el más similar a ése, en caso de enunciar un contrafáctico adecuado en ese mundo como para hacer referencia al nuestro.

Usemos un ejemplo estándar para ilustrar esto en palabras más sencillas. En nuestro mundo, en un local de ventas había una pérdida de gas combustible que llenó el ambiente de ese gas. Hubo una chispa en el tablero de electricidad y esa chispa ocasionó una explosión. Es fácil ver que sostendremos el siguiente condicional contrafáctico como verdadero: “de no haber ocurrido la chispa, la explosión no habría tenido lugar”. Esto puede sostenerse, según la semántica presentada por Lewis mediante el proceso siguiente. Se consideran todos los mundos posibles más similares al nuestro en el que la chispa no tuvo lugar. Llamaremos “mundo efectivo” al mundo que efectivamente tiene lugar, para distinguirlo de los demás mundos posibles y llamaremos “mundo base” al mundo en donde se enuncia el condicional contrafáctico. En nuestro ejemplo, el mundo base y el mundo efectivo pueden tomarse como *nuestro mundo*. Al realizar la evaluación del contrafáctico tenemos que tomar en cuenta, de entre todos los mundos posibles, un subconjunto que son los mundos accesibles, es decir dignos de ser tenidos en cuenta por el contexto. De todos ellos, al considerar los más similares al mundo base, no tendrían que diferir en casi nada que no fuera necesario para que la chispa no ocurriera. Así que son bien parecidos al nuestro.

En esos mundos en que no hay chispa y tampoco hay otros motivos para que tenga lugar la explosión, ya que en nuestro mundo solo hubo chispa y no otros factores asociados con la explosión, por lo cual, considerar mundos posibles en los que no hay chispa pero hay alguien que pase con un fósforo prendido o alguna otra de tales variantes, sería considerar mundos que no son suficientemente cercanos. Mundos que tuvieran esos agregados se ubican más lejos y por lo tanto no son dignos de tenerse en cuenta. De modo que en los mundos que debemos considerar no hubo chispa y por lo demás son coincidentes con el nuestro, aunque debemos chequear si en ellos tuvo lugar o no la explosión. Como podemos apreciar, al considerar los mundos más similares al nuestro en los que no hubo chispa, obtendremos que tampoco hubo explosión. Y de este modo llegamos a justificar que el contrafáctico en cuestión es verdadero.

Ahora bien, imaginemos que en uno de esos mundos en los que no hubo chispa ni explosión, digamos *w*, alguien pasa por el lugar, observa la pérdida de gas en el local y comenta:

“si hubiera habido una chispa, habría ocurrido una explosión”

² Para un análisis de este problema véase Díez (2014).

Para evaluar este contrafáctico, los lugareños de w , que ahora es el mundo base para ese contrafáctico, deben considerar los mundos más similares a w en los que sí tenga lugar una chispa en ese local. Nuestro mundo,³ o mundo efectivo, califica perfectamente para ese *casting*.

Nuestro mundo difiere en muy poco, solo en lo indispensable y que está relacionado con que aquí sí tuvo lugar la chispa y en w , no. Es decir, califica como uno de los más similares, si no el más similar. De modo que los habitantes de w al evaluar su contrafáctico, piensan en nuestro mundo, agregando una colección de mundos cercanamente similares, y concluyen que de haber habido alguna chispa, habría habido una explosión, tal como ocurrió aquí.

Así las cosas, la similaridad comparativa global, a pesar de no darnos garantía de simetría, está operando en los casos paradigmáticos como si efectivamente fuera simétrica.

Aparentemente es inocua la no simetría de la relación, ya que en los hechos referidos a través de los condicionales contrafácticos, en lo que importa, parece estar funcionando como simétrica.

Esto pudo haber sido el terreno aparentemente firme que llevó a Phil Dowe (2000, 2001) a sostener cierta relación en espejo entre situaciones de causación con omisiones. Como veremos, la causación con omisiones obliga a Dowe a apartarse muy levemente de su teoría de la causación de intercambio de cantidades conservadas y debe echar mano de condicionales contrafácticos. Es entonces cuando aparece de modo implícito el problema de evaluar si la similaridad es simétrica o no.

La teoría de Dowe (1992a, 1992b, 2000) no puede dar cuenta de los casos de causación que involucran omisiones del mismo modo que lo hace con eventos que efectivamente tienen lugar, ya que su propuesta consiste fundamentalmente en el intercambio de magnitudes físicas, o su conservación, para los casos en que existe un teorema de conservación. Por ejemplo, un choque de bolas de billar es una interacción causal y los procesos entrantes, la bola blanca dirigiéndose hacia la bola roja que está en reposo y la bola roja *estando* en reposo, son causas de los procesos salientes, la bola blanca desviada de su dirección original, y la bola roja moviéndose cuando anteriormente estaba en reposo. Por otra parte, la bola blanca dirigiéndose hacia la bola roja es un proceso causal, al igual que la bola roja que permanece en reposo. La cantidad de movimiento lineal, o momento, de la bola blanca en cada instante está causada por la cantidad de movimiento lineal que tenía anteriormente y de igual modo para la bola roja en reposo que conserva su cantidad de movimiento nula. La conservación de esa cantidad física a lo largo de ese proceso hace que sea un proceso causal. La ley de inercia es un tipo de ley de evolución causal en donde las condiciones previas causan las propiedades posteriores constituyendo casos paradigmáticos de procesos causales.

Así, la causación tiene lugar o bien en casos de interacción causal con intercambio de cantidades, o bien por medio de que algún objeto o proceso que es portador de cierta magnitud física que se conserva a lo largo de un proceso casual.⁴ Las interacciones causales y los procesos causales asociados a entidades o procesos portadores de magnitudes físicas para las que existe un teorema de conservación dan cuenta de todos los casos de causación, por lo cual todo otro caso que parezca ser un caso de causación pero no sea de alguno de estos dos tipos, no será una *causación genuina*.

³ La nomenclatura “nuestro mundo” o “mundo real” debe tomarse como indexical en la teoría de Lewis (1973). Véase Abeledo, Flichman, Miguel y Redondo (1992).

⁴ Existen varios problemas en identificar los portadores de estas cantidades. Véase Miguel y Paruelo (2003).

En los casos que involucran omisiones, aquello que no ha ocurrido no puede intercambiar alguna magnitud física con ninguna otra cosa que efectivamente haya ocurrido. Tampoco es posible que aquello que no ocurre sea portador de alguna cantidad física que se conserva a lo largo de cierto lapso. Las omisiones no pueden tenerse en cuenta ni en las interacciones causales ni en los procesos causales.

Recordemos que podrían darse casos con omisiones en donde la omisión juega el papel de causa, como por ejemplo la falla en el corte del termostato de la heladera causó que se congelaran las bebidas, o bien casos en los que la omisión juega el papel de efecto como por ejemplo el padre al detener a su niño que corría hacia la calle evitó el accidente (causó que no ocurriera), y otros en donde hay omisiones en ambos argumentos de la relación de causación pretendida, como por ejemplo, la sequía causó la deshidratación del ganado.

Debe mencionarse que Dowe tiene en cuenta esta dificultad desde el comienzo de su propuesta y por lo tanto su tratamiento para los casos con omisiones tiene un matiz particular. La propuesta de Dowe es que la causación que involucra omisiones no es una causación genuina sino una *quasi-causación*, y desde sus primeros trabajos la nombra con un asterisco (causación*) a la que nos dedicaremos en la siguiente sección.

Dowe (2000, p. 125) recuerda y enfatiza que hay intuiciones encontradas respecto a los casos de causación con omisiones. Por un lado hay casos en los que la intuición rechaza que se trate de causación genuina. Dowe llama a estos casos los de la “intuición de la diferencia” para resaltar que la intuición es capaz de reforzar y rescatar la diferencia entre la *causación genuina* y la *quasi-causación*. Por ejemplo, el padre omite detener a su niño que va corriendo hacia la calle, y sucede un accidente. Decir que el padre es causa del accidente porque falló en detener a su niño que corría hacia la calle parece un exceso, aun cuando es verdadero el contrafáctico “si el padre hubiera detenido a su niño, éste no habría sufrido un accidente”. Como es sabido, y Dowe lo resalta, estas diferencias pasan a primer plano en las discusiones sobre eutanasia en las que es crucial distinguir los casos de *matar* de los de *dejar morir*.⁵

Por otra parte, siguiendo a Dowe, están los casos que llama de la “intuición de ser lo mismo” (*intuition of sameness*). En estos casos la intuición da un veredicto según el cual es lo mismo que se haya obtenido el efecto mediante una cadena que involucra omisiones, que mediante una cadena de causación de eventos positivos. En esta categoría nos recuerda el ejemplo debido a Michael McDermott: decapitar a alguien es causa de su muerte, pero lo es en virtud de una prevención ya que cortar la cabeza previene los procesos que habrían causado que la persona continuara viva.

Entonces, ¿qué hacer con el problema de contar con intuiciones diferentes para casos diferentes?

Dowe pasa revista a una diversidad de salidas posibles y elige la propuesta de que la *quasi-causación* es causación posible, es decir causación en mundos muy cercanos al efectivo y con esta salida intentará mostrar que se justifica la existencia de intuiciones encontradas.

⁵ Véase McGrath (2003) y Miguel y Paruelo (2007).

2. La propuesta de Dowe

La propuesta de Dowe (2001) intenta dar cuenta con mayor detalle de la causación con omisiones frente a varias objeciones dirigidas hacia a la versión de Dowe (2000), aunque la definición formal es la misma en ambas versiones.

Como anticipamos, su propuesta consiste en una teoría contrafáctica de la quasi-causación.

Escribiremos de forma estándar las prevenciones y las omisiones de la siguiente manera.

- ‘A quasi-causó no-B’ (prevención)
- ‘no-A quasi-causó B’ (quasi-causación por omisión)

En donde A, B son eventos (o hechos) positivos genuinos.

Y para los casos de prevención por omisión (casos en donde ambos argumentos son omisiones) tenemos el análisis más sencillo:

- no-A quasi-causó no-B si, ni A ni B ocurrieron y
- si A hubiera ocurrido, A habría causado B

Donde A y B nombran eventos o hechos positivos.

En estos casos una omisión previene algo por el solo hecho de que cierto evento al no ocurrir no causó cierto otro evento entendido como efecto, y señala el ejemplo en que “el no tomar arsénico por parte de Juan, previno su muerte”.

Como se puede ver, este tipo de casos no son de causación genuina sino relacionados con una causación posible. El no tomar arsénico por parte de Juan previno su muerte es verdadero en el contexto en el que es verdadero que si Juan hubiera tomado arsénico, habría muerto, el cual es un enunciado contrafáctico acerca de un caso de causación genuina que no ocurrió en el mundo efectivo sino en mundos posibles lo más similares al mundo en el que se enuncia el contrafáctico.⁶ El mundo base es uno en el que Juan no tomó arsénico y no murió. Los mundos más similares al mundo base en donde Juan toma arsénico, son mundos donde Juan muere. Además, en esos mundos el arsénico causa la muerte por una serie de intercambios de magnitudes físicas que justifican identificar esos episodios como de causa y efecto. En esos mundos posibles la ingesta de arsénico causa genuinamente la muerte de Juan.

Los casos de prevención por omisión entonces, son casos que aluden a una causación genuina que podría haber ocurrido pero no ocurrió. Solo tiene lugar en los mundos más similares al mundo base en donde efectivamente tienen lugar los episodios que en el mundo base fueron omisiones.

Analicemos ahora los casos estándar en donde hay omisiones en uno solo de los argumentos de la relación, comenzando con los casos de prevenciones (P).

A *previno* B si A ocurrió y B no, y ocurrió cierto proceso o evento x tal que

(P1) hay una interacción causal entre A y el proceso debido a x, y

(P2) si A no hubiera ocurrido, x habría causado B.

Donde A y B son eventos o hechos positivos y x es una variable cuyo rango son eventos o hechos.

⁶ Los motivos para elegir la terminología “mundo efectivo” están en Abeledo y Flichman (1997).

Dowe aclara que la razón por la que se presenta el criterio únicamente como una condición suficiente, es porque existe otro tipo de prevenciones como los casos de sobredeterminación en la prevención que harían fallar el punto (P2), pero podemos dejarlos a un lado para esta discusión

Ilustra el criterio con el caso del padre que detiene al niño y así previene el accidente.

Hay una interacción causal entre A, la acción del padre, y el proceso x identificado con el niño corriendo hacia la calle, y es verdadero que si el padre no hubiera detenido al niño, el proceso x , el niño corriendo hacia la calle, habría sido causa del accidente.

Nos hace notar Dowe que existió una interacción causal genuina, aunque no entre A y B sino entre A y el proceso x .

Continuemos con los casos de quasi-causación por omisión.

Omisión: no-A quasi-causó B si B ocurrió pero A no, y ocurrió cierto x tal que

(O1) x causó B, y

(O2) si A hubiera ocurrido entonces A habría prevenido B al interactuar con x

En donde A y B son eventos o hechos positivos y x es una variable que recorre eventos o hechos y la prevención se analiza como más arriba.

Nuevamente ilustra el criterio. La falla del padre en detener al niño quasi-causó el accidente ya que el padre no detuvo al niño (omisión) y el accidente ocurrió (B) y había un proceso –el niño corriendo hacia la calle– tal que si el padre hubiera detenido al niño, eso habría prevenido el accidente.

Al igual que antes, advierte de la posibilidad de casos más complejos como casos de prevención por omisión en donde si A hubiera ocurrido, habría quasi-causado B y otras combinaciones como prevención por prevención y así siguiendo.

3. Ventajas y logros

Dowe enfatiza que su teoría sobre la quasi-causación no involucra una teoría particular de la causación y eso la hace más flexible.

Podemos decir que otra ventaja consiste en que tampoco involucra una teoría de contrafácticos. Estrictamente hablando podría utilizarse la teoría de Stalnaker (1968) o la del propio Lewis (1973) al que Dowe refiere sistemáticamente pero sin seguir paso a paso su teoría semántica.

La relación contrafáctica se establece entre eventos positivos evitando así el problema de lidiar con omisiones, problema que limita la aplicación de la semántica de Lewis para los condicionales contrafácticos.

También resuelve la intuición de Lewis de que algunas omisiones juegan el papel de causas, ya que en realidad estarían jugando el papel de quasi-causas, y no serían más que parientes cercanos de la genuina causación.

Dowe explica por qué surgen las intuiciones encontradas en tratar a veces las omisiones como causas genuinas y a veces rechazar ese trato. Propone que el origen de tal discrepancia radica en que, dado que la quasi-causación no es una causación genuina sino una afirmación contrafáctica sobre una causación posible, a los fines prácticos de la comunicación en muchas ocasiones esa diferencia no se tiene en cuenta, mientras que en otras ocasiones los hablantes quieren focalizarse sobre tal diferencia.

La situación en la que x causa y , es cercana a la situación en la que no ocurrieron x e y , pero de haber ocurrido x , habría causado y . El hecho posible no efectivizado permite tomar la ausencia de x como evidencia a favor de la explicación de por qué no ha ocurrido

y. Esta cercanía epistémica no significa que no valga la pena hacer la distinción conceptual para distinguir metafísicamente ambas situaciones.

Según Dowe, el hecho de que esa discrepancia no siempre ocupe el foco de atención, es atribuible a lo que llama “*epistemic blur*” y que podemos traducir por “borrosidad epistémica.”⁷

Dowe acepta que puede ocurrir que no siempre quede claro si estamos tratando con un evento o hecho positivo o con uno negativo. Los eventos que pensamos como negativos pueden resultar ser en realidad, positivos y viceversa. Hay entonces una borrosidad epistémica entre la quasi-causación y la causación genuina. Cita el ejemplo de Lewis en el que el opio parece causar genuinamente el adormecimiento, pero que el opio en realidad previene que algunos procesos normales realicen su función normal, y por tanto se trata en última instancia de un caso de quasi-causación y no de causación genuina como se entendía al principio.

Así, los casos de quasi-causación son mucho más comunes de lo que al principio podía reconocerse. Quizás el efecto de fumar sea prevenir ciertos procesos normales en ciertas células, y por tanto frente a la ausencia de tales procesos, la muerte celular prospera y por ese motivo sostenemos que el fumar causa enfermedades cardíacas, aunque en el fondo pueda tratarse de un caso de causación por omisión, es decir, un caso más de quasi-causación.

Dowe obtiene ventaja por sobre las corrientes que equiparan ambos tipos de casos y también supera a las corrientes que los separan de tal modo que no pueden dar cuenta de la equivalencia que tienen a nivel pragmático.

4. Dificultades remanentes

Los casos del estilo del ejemplo de la decapitación, sumados a la propuesta de Schaffer (2000) de causación por desconexión, nos dan un resultado nada alentador, aunque no concluyente, de que habría una mayor cantidad de casos de quasi-causation que los esperados, algo ya reconocido por Dowe respecto de una objeción similar,⁸ y por tanto, la quasi-causation sería tan o más habitual que la causación genuina.⁹ Este descubrimiento no es en sí mismo un obstáculo, hay que reconocerlo, porque podría haber muy pocos casos de causación genuina que nos dieran la ‘calibración’ de la teoría y luego con esta teoría causal poder dar cuenta de los casos de quasi-causation que refieren a causaciones genuinas en mundos posibles cercanos.

De todos modos, resulta extraño elaborar una teoría para muy pocos casos paradigmáticos cuyo gran éxito fuera su aplicación a casos de causación que no ocurren en el mundo efectivo sino en mundos alternativos.

⁷ Dowe (2004) p. 194.

⁸ Esta objeción debe distinguirse de la de Beebe que apunta a la causación indirecta por estar ahora posibilitada una cadena con peldaños de omisiones. Dowe se refiere un artículo inédito de Beebe: Causes, omissions and conditions. Sin embargo, Beebe (en comunicación personal 2016) comenta que la parte citada por Dowe fue incluida en Beebe (2004). El ejemplo de Beebe se refiere a que si alguien no se limpia la oreja, su omisión es quasicausa de una inundación dentro de 4 años. En Beebe (2004) se hace claro que la historia causal del mundo permite elegir explicaciones y que en ocasiones pragmáticamente podemos elegir explicaciones que echan mano de ausencia de factores que habrían impedido que ciertos eventos no ocurrieran. Dowe agrega que muchos de estos problemas de causación indirecta provienen de la dificultad de aceptar la transitividad causal, y que no es un problema particular de su propuesta sino de cualquier teoría de la causación. Véase el problema de la transitividad causal en Miguel (2007).

⁹ En Schaffer (2004) aparecen otras críticas externas a la teoría por lo cual no las retomamos aquí.

Por otra parte tal situación se ajusta más bien a un análisis conceptual de por qué los hablantes llaman causación a estos casos de quasi-causation en vez de uno empírico como el que Dowe persigue como objetivo central.

Sin embargo una dificultad mayor se vislumbra en el criterio de quasi-causación por omisión.

Recordemos que A no ocurrió y un proceso x causó B, y que si hubiera ocurrido A, habría prevenido B al interactuar con x, interrumpiéndolo.

Vemos que esta definición a su vez depende de la de prevención que consiste en otro contrafáctico, A previno B, si de no haber ocurrido A, habría ocurrido B en virtud de un proceso x que A interrumpió interactuando causalmente (ver Figura 1).

Entonces, no-A quasi-causea B (fig. 1. derecha) si es cierto que de haber ocurrido A, habría prevenido B (fig. 1. izquierda). Y a su vez, A previene B (fig. 1. izquierda), si es cierto que de no haber ocurrido A, x habría causado B (fig. 1. derecha). El mismo Dowe (2000) sostiene que se da una situación en espejo entre las omisiones y las preventiones.

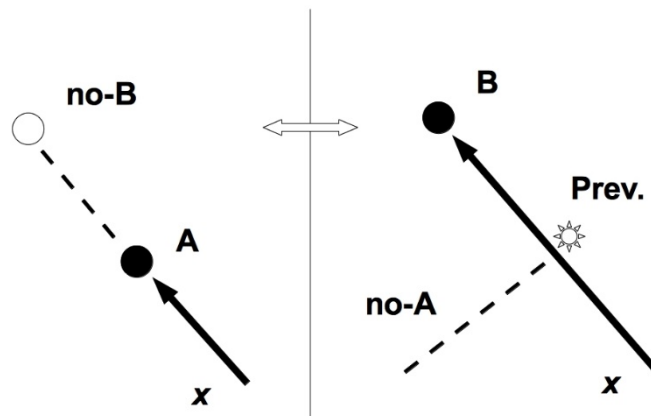


Figura 1. Izquierda: prevención. Derecha: omisión de prevención.

Aunque comienza diciendo que eso ocurre en las circunstancias adecuadas, no especifica a qué se refiere y continúa su análisis de modo general. Muestra la relación en espejo del siguiente modo. Cuando un preventor A causó* no-B, si no hubiera ocurrido A, la omisión no-A habría causado* B por omisión; y donde la omisión no-A causó* B, si A hubiera ocurrido, habría prevenido B.¹⁰

Pero esto debe decirse de forma indexada a los mundos en donde se hacen esas afirmaciones. Así que deberíamos indicar en qué mundos ocurre cada cosa, incluida la

¹⁰ Dowe (2000, p. 138): “In fact, in the right circumstances, the following mirror relation holds between prevention and omission: Where a preventer A caused* not-B, had A not occurred, not-A would have caused B by omission; and where an omission not-A caused* B, had A occurred, A would have prevented B.” En esa misma página: “In both cases, the scenario we imagine in the all-important counterfactual is just the mirror case.”

enunciación de cada frase. En pos de la simplicidad del análisis, dejaremos a un lado el problema del supuesto de límite y nos referiremos al mundo más cercano en vez de a los mundos más cercanos. El éxito del análisis no depende de esta simplificación aunque su presentación se hace más comprensible. La redacción debe quedar del siguiente modo.

Cuando un preventor A causó* no-B, eso indica que en nuestro mundo M_E (mundo efectivo) ocurrió A y no ocurrió B.

En ese mismo mundo base M_E , enunciamos que “si no hubiera ocurrido A, la omisión no-A habría causado* B por omisión” con eso hacemos referencia a otros mundos muy similares al mundo M_E pero en los cuales A no ha ocurrido y queremos decir que en esos mundos B habría ocurrido. Digamos que esos mundos son mundos M_1 .

De modo similar, si en nuestro mundo M_E , A no ocurrió y B sí, diremos que la omisión no-A causó* B, y este mismo mundo M_E afirmaremos que si A hubiera ocurrido, habría prevenido B. Ahora bien, en la forma de analizarlo que hemos llevado adelante, en ambas ocasiones hemos elegido que sea el mismo mundo base nuestro mundo efectivo. Pero no hemos analizado que tanto las relaciones “...causó*...” como la relación “...previno...” tienen embebidas definiciones que hacen referencia a otros mundos y por lo tanto el análisis deberá ser más fino todavía.

Por ahora vemos que los preventores que ocurren causan* omisiones y las omisiones, de haber ocurrido habrían sido preventores.

El punto que evita que el criterio sea circular es que el último peldaño tanto en las que las omisiones juegan el papel de causa como cuando se trata de preventores es que alude a que x habría causado B.

Ahora debemos indexar las situaciones respecto de los mundos posibles no solo en lo que refiere a si el preventor no hubiera ocurrido, sino a su carácter de preventor y, de modo similar con las omisiones.

Veamos entonces cuál es el mínimo de mundos involucrados en la presunta relación en espejo.

1. no-A quasi-causa B (fig. 1. Derecha) en M_E (mundo efectivo).
2. A cuasi-causa no-B (A previene B) en M_1 (mundo más similar a M_E).
3. A interactúa con x interrumpiéndolo en M_1 (mundo más similar a M_E).
4. A previene B en M_1 ya que en el mundo M_2 más cercano a M_1 en donde A no ocurre, el proceso x no se interrumpe y efectivamente causa B.

Vemos que no hay una verdadera situación en espejo si no garantizamos la simetría en la similaridad de mundos posibles, es decir que si M_1 es el mundo más similar a M_E y M_2 es el mundo más similar a M_1 entonces M_2 debe ser el mismo que M_E para afirmar que hay una situación en espejo.

Hasta aquí el análisis general. Tomemos ahora un caso que provea una intuición adecuada y firme para comprender las dificultades en asegurar que en la práctica se pueda asumir la simetría.

En lo que sigue se muestra un ejemplo en el que el mundo M_2 no es el mismo que M_E y en ese mundo no ocurre que x cause B, aun cuando cumple en ser el mundo más cercano en el que no ocurre A, y también allí se da el proceso x .

El ejemplo se compone de los siguientes elementos: en el mundo efectivo M_E , no-A representa la ausencia de un dique que proteja a la ciudad, x representa la crecida del río, y B representa la inundación de la ciudad. Cada vez que llueve mucho, crece el río y la ciudad se inunda. Los habitantes están convencidos de la verdad del contrafáctico según el cual “si hubiera habido un dique, la ciudad no se habría inundado” que corresponde perfectamente a las condiciones de ausencia de un preventor.

El mundo M_1 más cercano al mundo efectivo en donde hay un dique es un mundo en donde efectivamente A previene la inundación B. Hasta allí todo sucede como imagina Dowe.

Pero una vez ubicados en M_1 en donde el dique previene las inundaciones, debemos evaluar el contrafáctico “si A no hubiera ocurrido, entonces x habría causado B” es decir, si no hubiera estado el dique, la crecida habría inundado a la ciudad.

Este contrafáctico parece correcto, pero el mundo M_2 más cercano al M_1 que es necesario tomar en cuenta para su evaluación, podría no ser un mundo simplemente sin dique que coincidiera con el mundo efectivo que no tiene ninguna construcción en absoluto, sino que podría ser uno en el que hay un dique en construcción y sin terminar. Es decir que en un mundo M_1 en el que hay un dique, el contrafáctico nos lleva a que el mundo más cercano es M_2 , uno en el que el dique no está completo y no a M_E en el que todavía ni siquiera ha comenzado a construirse. La gradualidad con la que desaparece el dique en la suposición contrafáctica va acompañada de una gradualidad en la importancia de la inundación. Este es un primer motivo por el cual no hay simetría entre omisiones y preventores.

En el fondo estamos tratando de reflejar en el ejemplo la topología de ordenamiento de mundos que mostramos en la Figura 2. Según la figura la similitud aplicada de modo anidada en un contrafáctico no necesariamente da un resultado simétrico.

Alguien podría sostener que al enunciar en M_1 el contrafáctico que se refiere a la ausencia del dique los mundos más similares son aquellos en los que no existe nada en absoluto y entonces este argumento parecería no cumplir el cometido de mostrar la falla de la simetría en la similaridad de mundos. Por lo cual pasaremos a un segundo tipo de problema que introduce obstáculos insalvables para sostener la simetría.

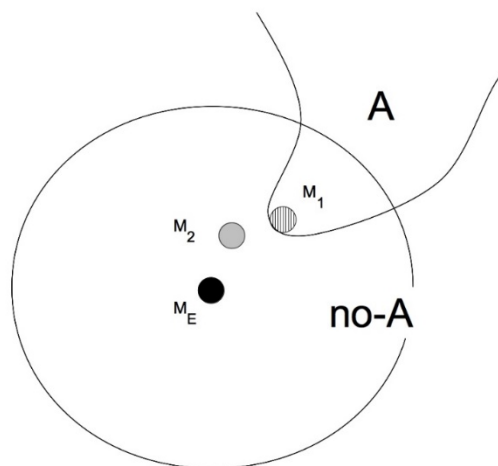


Figura 2. Ordenamiento de mundos posibles. Para los contrafácticos enunciados en M_E se considera M_1 como el mundo más similar para la evaluación. Pero para cuando se enuncia contrafácticos desde M_1 se debe considerar a M_2 como el más cercano a M_1 y no a M_E como la percepción simétrica indicaría.

El segundo ejemplo es como sigue. Oscar es un joven cuyas piernas están amputadas pero ama el deporte y ha logrado dedicarse a las carreras pedestres gracias a

unas prótesis que suplen las piernas, le dan gran velocidad y agilidad. Oscar comenzó corriendo en carreras para personas con capacidades diferentes. Luego de ser campeón en esa categoría pidió correr disputando el título en una competencia sin hándicap contra corredores que no tienen ninguna característica especial en su anatomía. Se ha preparado una carrera en camino de montaña entre Oscar y Juan, que es campeón mundial sin ninguna amputación.

Juan se ha entrenado mucho más que Oscar para esta carrera por lo cual su estado físico es mejor que el de Oscar. Al iniciarse la carrera, Oscar sale con mayor ímpetu y unos metros atrás va Juan, que confía poder ganarle gracias a su mejor preparación para una carrera tan larga.

En un tramo de la carrera, unas piedras se desprenden de la montaña y caen sobre Juan, que queda un poco lastimado pero sigue corriendo. Las piedras no alcanzaron a Oscar que iba unos metros adelante en ese tramo.

Finalmente Oscar gana la carrera. Sus prótesis parecen ser el motivo de su superioridad. Al terminar la carrera el entrenador de Juan comenta “si Juan hubiera usado esas prótesis, habría ganado la carrera” y agrega “si Juan hubiera usado esas prótesis, no habría recibido el impacto de las piedras, ya que habría ido unos metros adelante, como lo hacía Oscar en ese tramo”. Todo esto indicaría que Juan está en mejor estado para ganar la carrera pero que la diferencia se debió al uso de esas prótesis que brindan una mayor *performance* a quien las use.

La ausencia de prótesis en Juan parece ser la causa por omisión de que recibiera las piedras.

En términos del análisis de Dowe, la omisión de prótesis, no-A, causa* B, que Juan reciba las piedras, y todo esto sucede en el mundo efectivo M_E . Las prótesis habrían cumplido la función de prevenir que Juan recibiera las piedras que lo dañaron.

Cuando el entrenador enuncia su contrafáctico, se debe considerar un mundo M_1 en el que Juan usa prótesis. Pero esta característica también nos hace considerar que Juan en ese mundo tiene las piernas cortadas. Es decir que, en M_1 Juan está en las mismas condiciones anatómicas que Oscar pero sigue estando mejor preparado físicamente. Podría pensarse que en ese mundo Juan solamente usa las prótesis pero tiene las piernas igual que en el mundo efectivo. Pero esto tiene varias dificultades. Por una parte, son prótesis y no agregados. La sola noción de prótesis indica que suple una parte del cuerpo que no está. Por otra parte, es muy probable que si Juan se colocara esas mismas prótesis con las piernas que tiene en el mundo efectivo, no podría ni dar un paso hasta tanto se entrenara y, claro, en caso de que hubiera tenido las piernas amputadas y hubiera usado esas prótesis habría tenido que ser su forma habitual de correr y entrenarse, algo que no habría ocurrido si conservaba las piernas. Se podría agregar que la organización de la competencia permite las prótesis porque suplen las piernas y no como un adicional que pueda elegir un corredor con piernas.¹¹

Aclarado este punto vemos que el mundo M_1 se aparta del mundo efectivo en varios aspectos y no es posible alejarse menos, tal como suele ocurrir para algunos condicionales contrafácticos cuya suposición necesita de varios cambios simultáneos para cumplir con la cláusula contrafáctica. En M_1 Juan usa prótesis, Juan tiene sus piernas amputadas y sabe usar esas prótesis. De otro modo no usaría prótesis y de no haberse entrenado con ellas, no habría sido un rival para Oscar. Ese es nuestro mundo más cercano para evaluar el contrafáctico enunciado por el entrenador.

¹¹ Agradezco a Diego Letzen su intervención en ocasión de exponer una versión anterior de este trabajo. Sus comentarios me permitieron prestar atención a este punto.

Ahora bien, veamos qué sucedería si en ese mundo M_1 alguien evaluara el contrafáctico “si Juan no hubiera usado las prótesis, habría recibido las piedras al pasar por la montaña”. Este contrafáctico está enunciado en M_1 y debemos buscar cuáles son los mundos más cercanos en los que Juan no usa sus prótesis. Ese mundo es uno en el que Juan, con sus piernas amputadas no puede correr ni siquiera a la velocidad de su contraparte: Juan con piernas completas. Juan directamente no podría ganar esa carrera, pero además, habría ido tan lentamente que no habría pasado por la zona de las piedras cuando éstas caían sino que habría pasado más tarde. Al enunciar ese contrafáctico en M_1 deberemos considerar otro mundo M_2 que sería el más similar a M_1 pero en donde Juan sigue teniendo las piernas amputadas pero no usa prótesis. Este mundo no coincide con el mundo efectivo inicial. No es posible sostener una relación en espejo y la simetría queda absolutamente vedada incluso desde una clara intuición y no solamente desde el punto de vista formal.

Como puede apreciarse la construcción del ejemplo es muy sencilla, se deben buscar casos en donde sea necesario hacer al menos dos cambios para cumplir con la cláusula contrafáctica. Las clásicas e incipientes consideraciones de Nelson Goodman (1947) repican una y otra vez en el desarrollo de las discusiones sobre contrafácticos, leyes y explicaciones. Goodman nos advertía acerca de una serie de condiciones que tenían que ser atendidas al evaluar el contrafáctico. Entre ellas estaba el problema de la cosostenibilidad de enunciados que tenían que acompañar la suposición contrafáctica para no suponer un conjunto contradictorio de premisas. Con solo recordar sus advertencias se infiere que es posible construir ejemplos en los que la relación en espejo quede desafiada. El ejemplo que hemos presentado muestra que existen casos prácticos y con intuiciones claras para los que, dado que hay más de un cambio que realizar para la suposición contrafáctica, no se puede asumir una equivalencia pragmática y no hay posibilidad de suponer una simetría en la similaridad comparativa entre mundos posibles.

Se puede ahora distinguir entre dos tipos de contrafácticos. Una clase de contrafácticos cuyo antecedente requiere solamente una modificación ligada a la ocurrencia del consecuente y por lo tanto solo una condición relacionada con la evaluación del condicional. En estos casos la modificación necesaria para ubicar el mundo posible M_1 más similar al mundo base puede ser exactamente la inversa que la que se necesita para ubicar al mundo más cercano a M_1 y dar como resultado una simetría en la similaridad de mundos que favorece la equivalencia en tratar los casos de causación con omisiones como casos genuinos, de acuerdo a la *equivalencia pragmática* que sostiene Dowe.

Otro tipo de contrafácticos, en cambio, hace indispensable suponer varias modificaciones (al menos dos) para encontrar al mundo M_1 más similar al mundo base M_E de modo que luego, al evaluar contrafácticos que se enuncian en ese mundo M_1 las suposiciones mínimas pueden ser menos que las asumidas para el primer movimiento y de ese modo nos llevan a un mundo M_2 más similar al M_1 y no coincidente con M_E .

La similaridad comparativa global entre mundos posibles es un viaje de ida. Es por ese motivo que la evaluación de contrafácticos anidados debe ser tratada con mucha mayor delicadeza de análisis.

Me habría gustado concluir este trabajo reflexionando acerca de que si hubiéramos tenido en cuenta esta dificultad, no habríamos utilizado con tanta confianza las teorías de la causación para las que las omisiones podrían haberse tratado como eventos que de haber ocurrido habrían causado algún efecto.

5. Referencias

- Abeledo, H. & Flichman, E. H. (1997). "Mundo efectivo" o "mundo real" o actualmente se traiciona así. *Crítica, Revista Hispanoamericana de Filosofía*, 29(87), 101-116.
- Abeledo, H. & Flichman, E. H., Miguel, H., Redondo, C., (1992). Descentración, indexicalidad y compromiso ontológico. *Revista Latinoamericana de Filosofía*, XVIII(2), 241-258.
- Beebee, H. (2004). Causing and Nothingness. En L. A. Paul, E. J. Hall & J. Collins (eds.), *Causation and Counterfactuals*. The MIT Press 291-308.
- Díez, J. A. (2014). Contrafácticos, condicionales estrictos, accesibilidad y supuesto de límite. En H. Miguel (ed.), *Causación, explicación, y contrafácticos*; 245-278. Buenos Aires, Argentina: Editorial Prometeo.
- Dowe, P. (1992a). Process Causality and Asymmetry, *Erkenntnis*, 37, 179-196.
- Dowe, P. (1992b). Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory. *Philosophy of Science*, 59, 195-216.
- Dowe, P. (2000). *Physical Causation*. New York, USA: Cambridge University Press.
- Dowe, P. (2001). A Counterfactual Theory of Prevention and 'Causation' by Omission. *Australasian Journal of Philosophy*, 79(2), 216-26.
- Dowe, P. (2004). Causes are Physically Connected to their Effects: Why Preventers and Omissions are not Causes. En Christopher Hitchcock (Ed.). (2004), *Contemporary Debates in Philosophy of Science*. Oxford, UK: Blackwell Publishing Ltd. 189-196.
- Goodman, N. (1947). The Problem of Counterfactual Conditionals. *The Journal of Philosophy* XLIV, 113-28.
- Hitchcock, C. (Ed.). (2004). *Contemporary Debates in Philosophy of Science*. Oxford, UK: Blackwell Publishing Ltd.
- Lewis D. (1973). *Counterfactuals*. Oxford, USA: Blackwell.
- McGrath, S. (2003). Causation and the Making/Allowing Distinction. *Philosophical Studies* 114(1-2), 81-106.
- Miguel, H. & Paruelo, J. (2003). Overlapping Causal Interactions in Phil Dowe's Theory. *Análisis Filosófico*, XXII(1), 69-84.
- Miguel, H. & Paruelo, J., (2007). Causar o dejar que ocurra. *Andamios, Revista de Investigación Social*, 4(7), 7-18.

- Miguel, H., (2007). La transitividad causal a escena. En Luis Salvático y Pío García (eds.), *Epistemología e Historia de la Ciencia. Selección de las XVII Jornadas*, 13, 369-375.
- Miguel, H., (Ed.). (2014). *Causación, explicación, y contrafácticos*. Buenos Aires, Argentina: Editorial Prometeo .
- Schaffer, J. (2000). Causation by disconnection. *Philosophy of Science*, 67, 285–300.
- Schaffer, J. (2004). Causes need not be Physically Connected to their Effects: The Case for Negative Causation. En C. Hitchcock (Ed.), *Contemporary Debates in Philosophy of Science*. Cornwall, UK: Blackwell.
- Stalnaker, R. (1968). A Theory of Conditionals. En Nicholas Rescher (ed.) *Studies in Logical Theory*. Oxford: Blackwell, American Philosophical Quarterly, Monograph Series. 2: 98-112. [Reimpreso en Ernest Sosa (ed.) 1975: 165-179].

Modelos y representación

Hernán Lucas Accorinti¹
Juan Camilo Martínez González²

Recibido: 28 de diciembre de 2015.

Aceptado: 26 de julio de 2016.

Resumen. Según la epistemología tradicional, los modelos devienen epistémicamente relevantes en la medida en que logren *representar* al fenómeno. Desde esta perspectiva se intentó justificar tal relación a partir de un *isomorfismo* o una *similitud* entre ambos dominios. En el presente trabajo evidenciaremos, a partir de ciertos tipos de modelos utilizados comúnmente en el ámbito científico, que una postura que pretenda naturalizar la representación excluyendo al agente como vínculo entre el modelo y el sistema será siempre inadecuada. A su vez, sobre la base del análisis de dos modelos incompatibles propios de la química cuántica como son el *Enlace de Valencia* y el *Orbital Molecular*, explicaremos en qué sentidos podemos afirmar que la potencialidad representativa no es una condición necesaria para la construcción del conocimiento científico. En efecto, en química cuántica el conocimiento se genera a partir de la capacidad de cálculo y predicción de los modelos; cálculos y predicciones no siempre susceptibles de ser comprendidos desde marcos epistemológicos representacionistas.

Palabras clave: modelos – representación – química cuántica – modelos moleculares.

Title: Models and representation.

Abstract. According to traditional philosophy of science, models become epistemically relevant insofar they are able to *represent* the phenomenon. From this perspective, it has been attempted to justify such a relation in terms of an *isomorphism* or *similarity* between both domains. The present work will show, on the basis of some kinds of models commonly used in science, that any stance that aims at naturalizing representation by excluding the agent as a mediator between model and target system will always be insufficient. In turn, by analyzing two incompatibles models of quantum chemistry such as *Valence Bond* and *Molecular Orbital*, we will show that representative potentiality is not a necessary condition for the construction of scientific knowledge. Indeed in quantum chemistry knowledge is the result of the computing and prediction power of the models themselves; computing and predictions that are not always susceptible to be understood from representationalist epistemological frameworks.

Keywords: models – representation – quantum chemistry – molecular models.

¹ Universidad de Buenos Aires

✉ hernanaccorinti@gmail.com

² CONICET – Universidad de Buenos Aires

✉ olimac62@hotmail.com

Accorinti, Hernán Lucas; Martínez González, Juan Camilo (2016). Modelos y representación. *Epistemología e Historia de la Ciencia*; 1(1); 21-34. ISSN: 2525-1198



1. Introducción

En el presente trabajo analizaremos el problema de la representación tal como se presenta en el contexto de algunos modelos que son utilizados en la actualidad por la práctica científica. La especificidad y particularidad del tema se debe a que la potencialidad representativa de los modelos en ciencia excede el ámbito del lenguaje, autonomizándose así del clásico problema en torno a la referencialidad de éste. En efecto, si interpretamos a los modelos como entidades no estrictamente lingüísticas, nos enfrentamos al particular dilema de saber cómo puede ser que un objeto, y no un artefacto lingüístico, represente a otro objeto.

La importancia de la pregunta por la capacidad representativa de los modelos se debe a que, según la posición tradicional, los mismos devienen epistémicamente relevantes en la medida en que *representan* al fenómeno. Esta perspectiva asume que los modelos nos dicen algo acerca de la realidad si los aspectos presentados en ellos encuentran su contrapartida en el objeto de estudio (Frigg & Hartmann, 2006). En otras palabras, se supone que los modelos y los *sistemas target* (esto es, aquella parte o aspecto del mundo del cual se ocupa el modelo)³ tienen una determinada relación, que en un principio pretendió ser no intencional, en virtud de la cual se fundamenta la potencialidad cognoscitiva de los modelos. Dicha relación es analizada o bien a partir de algún tipo de morfismo en términos de una relación matemática definida extensionalmente a partir de una correspondencia entre estructuras, o bien en términos de una similitud, en donde, estableciendo una relación más débil que las propuestas estructuralistas, remite a una mera semejanza respecto a algunos aspectos entre dos dominios.

Ahora bien, es evidente que los modelos no son idénticos al *sistema* ya que, en tanto pretenden ser *re-presentaciones* y no meras presentaciones, se instituyen como un sustituto y, por lo tanto, como algo diferente de aquello a representar. Este tipo de consideraciones refieren al hecho de que todo modelo se constituye, por un lado, como producto de un proceso de abstracción de aquellas variables no relevantes del *sistema* y, por el otro, como producto de un proceso de idealización de aquellos factores que, no encontrándose en el *sistema*, son fundamentales para la manipulabilidad y, con ello, para la eficacia del modelo. En este sentido, la crítica a la perspectiva representacionista no se debe a que se espere de ella una imagen especular. Aun la postura más rigurosa como el *isomorfismo*, que requiere que todos los elementos de un conjunto (modelo) tengan una y sólo una relación con los elementos de otro conjunto (*sistema*)⁴, considera tales aspectos.

Sin embargo, asumir que los modelos pueden decirnos *al menos algo* respecto de la realidad en términos representacionales, lejos de agotar la discusión, abre un campo de análisis sumamente fructífero respecto de lo que dicha postura implica.

En relación con ello, en primer lugar evaluaremos la perspectiva representacionista (tal como se la concibió originalmente en términos no intencionales) a partir del análisis de ciertos tipos de modelos comúnmente utilizados en ciencias. Es así que en el punto 2 analizaremos los modelos a escala; en el punto 3 los modelos de datos; en el punto 4 los modelos matemáticos y en el punto 5 el problema generado por

³ De aquí en adelante cuando hablemos de sistema nos referiremos a lo que en la literatura se conoce como “sistema target”.

⁴ Adoptaremos, así, la definición brindada por Suárez según el cual el isomorfismo es una función de uno a uno que vincula todos los elementos de un dominio con los elementos de otro dominio y viceversa (Suárez, 2003, p. 228).

los procesos de idealización. Cada uno de estos apartados contribuirá a evidenciar que el lazo establecido por el imaginario cultural entre conocimiento y representación, a pesar de ser vital para cualquier tradición realista, no siempre encuentra suficiente fundamentación filosófica frente a la multiplicidad de modelos existentes en la práctica científica. En efecto, como explicaremos ulteriormente, el principal problema de la perspectiva representacionista fue que en un principio se pretendió fundamentar tal relación en términos no intencionales en virtud a ciertas características *propias* compartidas entre modelo y sistema.

En continuidad con tales propósitos, en segundo lugar, sobre la base del análisis de un caso específico de la química cuántica, en particular, el caso de los modelos de *Enlace de Valencia* y de *Orbital Molecular*, radicalizaremos las críticas afirmando que existen buenas razones para pensar que los modelos, tal como se desarrollan en algunas disciplinas científicas, pueden brindar información respecto al mundo aun cuando ésta no pueda ser evaluada en términos representativos.

Por último, en tercer lugar, analizaremos la posibilidad de revalidar las perspectivas representacionistas desde una postura pragmatista que no naturalice el proceso de la representación. En este sentido afirmaremos que, a pesar de que el giro pragmático se instituye como francamente superador en tanto que introduce el papel del agente, ello lo hace a costa de deflacionar radicalmente el concepto mismo de representación.

2. Modelos a escala

Comenzaremos analizando los modelos a escala intentando caracterizar los parámetros bajo los cuales los mismos pueden constituirse como instancias representativas. La relevancia de este tipo específico de modelos es que han sido uno de los primeros modelos utilizados y, si bien han perdido cierto peso a partir de los modelos computacionales, aún siguen siendo utilizados en áreas como la ingeniería o la arquitectura.

Los modelos a escala se caracterizan por ser entidades u objetos físicos concretos, útiles para predecir, describir e investigar el comportamiento de otros sistemas físicos. En este sentido, estos modelos se constituirían a partir de una analogía entre dos sistemas físicos. Un caso paradigmático de los modelos a escala es el del “túnel de viento”. En él se intenta simular los efectos del aire en objetos como aviones, automóviles o misiles a partir de un conducto en el que se propulsa aire en forma constante. Como es de suponer, una de las dificultades iniciales, aunque relativamente menores, en este tipo de modelos es considerar las cuestiones relativas al tamaño y la proporción de los elementos empleados, como también a los materiales utilizados, ya que éstos pueden alterar el proceso experimental. Para el caso del “túnel del viento”, por ejemplo, no es menor mantener una proporción adecuada entre el tamaño y la velocidad del viento. En este sentido, la expectativa de construir modelos a escala lo más realista posibles supone desafíos técnicos y económicos que otro tipo de modelo carece.

Ahora bien, si pensamos que los modelos brindan conocimiento *porque* representan al fenómeno en cuestión, parecería natural suponer que, en tanto el modelo a escala es una *réplica* del original, lo que legitimaría el conocimiento establecido sería la *similitud* relativa a algunos aspectos existentes entre ambos dominios. Como afirma Mauricio Suárez (2003, p. 15), debido a que los dominios relacionados son entidades físicas concretas que establecen la relación a partir de su ser objeto, y no a partir de la

estructura que ejemplifican, no parecería intuitivo atribuir de un modo directo al *isomorfismo* (entendido como relación matemática biyectiva entre dos dominios) la potencialidad representativa.

Aparentemente, el isomorfismo cuenta con una ventaja sobre la similitud, ventaja que deriva del supuesto de que la representación se constituye en virtud de una relación natural no ambigua entre la estructura del modelo y la estructura del mundo. Sin embargo, esta ventaja se desvanece si se considera que entre un modelo y su sistema pueden establecerse diferentes tipos de relaciones estructurales. En este sentido, la estructura del fenómeno de estudio no está “allí” aguardando a ser descubierta, sino que depende de los intereses del científico, quien puede destacar diferentes relaciones que se establecen entre los diferentes elementos que componen al sistema. Por otro lado, otra desventaja del isomorfismo es que, a diferencia de la similitud, es incapaz de tomar en cuenta significativamente la diferencia de diversos modelos a escala del mismo sistema, debido a que, en tanto isomórficos, todos los modelos serían equivalentes.

Ahora bien, lo que el análisis pone de manifiesto es que, aun cuando la similitud aplica más adecuadamente para el caso de los modelos a escala, dista de ser una explicación completamente satisfactoria. En tanto que un objeto puede ser similar a otro en un campo infinito de aspectos no relevantes, la relación entre el modelo y el sistema no puede establecerse exclusivamente, como supone la versión naturalista de la representación, a partir de las propiedades intrínsecas a ellos. El modelo, incapaz por sí mismo de seleccionar el rasgo respecto del cual se dice similar al sistema, requiere un aditamento que consigne la propiedad compartida que determina la representación. Uno de los problemas que emergen de tal consideración es que, si la representación depende de la similitud relevante o elegida, y ésta, a su vez, es tal en función a que permite la representación, la caracterización buscada es indefectiblemente circular ya que estaríamos determinando la representación en función de una similitud que se torna preeminente por el hecho de ser representativa (Suárez, 2003, p. 26). A su vez, considerando que, estrictamente, un objeto siempre es similar en algún aspecto a otro objeto, la similitud se podría establecer tanto como condición de la representación como de la no representación.

Por lo tanto, en lo que respecta a los modelos a escala, la similitud, que se constituiría como la propiedad principal en virtud de la cual este tipo de modelos podrían llegar a ser representativos, no logra establecer las bases fundamentales como para legitimar la pretendida capacidad representativa.

3. Modelos de datos

Los modelos de datos cobran relevancia respecto del problema de la representación debido a que constituyen la primera instancia en el proceso representacional. Como afirmara Patrick Suppes (1962), las teorías no se contrastan con lo que suele denominarse “datos crudos”, obtenidos de la experimentación directa, sino con modelos de datos. Considerando esto el análisis de este tipo de modelos deviene imprescindible para el tema que nos convoca.

En este sentido, la base empírica de la teoría viene conformada por datos ya modelizados que, producto de las consideraciones pragmáticas y teóricas de los investigadores, no pueden considerarse como una mera copia de la realidad. A su vez, las partes del mundo que se pretenden representar no están estructuradas en un sentido obvio e inmediato. Suponiendo por un momento que es posible adscribir estructura a lo

real, el mundo no sería capturado inmediatamente sino que se presentaría modelizado de algún modo; y, por lo tanto, estrictamente hablando, el isomorfismo requerido por algunas posturas representacionistas se convertiría en una relación entre un modelo teórico y un modelo empírico (Knuuttila, 2005).

Evidentemente, el problema surge porque los modelos de datos, entendidos como rectificaciones, correcciones o selecciones de los datos provenientes de la experimentación, exigen algún tipo de manipulación e involucran supuestos en el proceso de su interpretación. Dependiendo de los elementos relegados y destacados, y del lazo establecido entre ellos, se producen diferentes modelos de datos. En efecto, la estrategia usual es introducir cambios en los conjuntos de datos, ya sea al eliminar los elementos presuntamente irrelevantes o los que resultan de errores de medición, o al establecer la relación entre los puntos seleccionados para la conformación del diagrama. Dependiendo de los cambios introducidos, se constituirán diferentes modelos. Cómo decidir respecto de los datos a tener en cuenta, y qué curva o lazo elegir dado un conjunto de datos, está condicionado por los objetivos teóricos y pragmáticos de la investigación. Esto pone en cuestión la capacidad representativa de los modelos de datos en términos de propiedades “naturales” de los dominios involucrados. A su vez la determinación de los denominados “datos crudos” supone un contexto interpretativo particular, que se constituye como parte del dominio de decisiones parciales que los científicos deben tomar (Harris, 2003).

Podría considerarse como ejemplo de un modelo de datos el resultado de la línea imaginaria trazada a partir de ciertos puntos pertenecientes a la trayectoria planetaria. Ahora bien, para obtener tal trayectoria se introdujeron cambios debido tanto a los datos que fueron descartados como a la configuración de un continuo a partir de una secuencia discreta de puntos. Por ello, podríamos preguntarnos por las entidades básicas que constituyen el dominio de la estructura, puesto que la estructura del sistema a representar varía en función de los objetos seleccionados como tales y al recorte parcial del todo compacto que tal selección impone. Consecuentemente, el mundo no tendría una estructura en estado bruto y, por lo tanto, el isomorfismo no podría establecerse como criterio natural y objetivo de representación. Desde esta perspectiva, la mentada trayectoria planetaria no contendría una estructura *per se* que permitiera hablar en términos de isomorfismo. Para ilustrar lo antedicho, Roman Frigg (2006) apela al ejemplo de la molécula de metano donde, según los objetos que constituyan el dominio (ya sea que se considere como objetos a los vértices o a las líneas que conforman la figura geométrica), se vislumbrarán diferentes estructuras que, aun teniendo el mismo sistema geométrico como base, no son isomórficas. Esto se debe a que tales estructuras poseen distinta cardinalidad: en virtud de la elección varían tanto el número de objetos como la cantidad de relaciones.

Por lo tanto, el modelo de datos, si bien es la instancia sobre la que se fundamenta toda pretensión representativa, lejos de ser aporofémico, resulta incapaz para justificar una concepción naturalista de la representación. En este sentido si se quisiera justificar la representación en términos estructuralistas mediante el modelo de datos habría primeramente que considerar que la mentada relación entre modelo y sistema no se produce de un modo inmediato sino que la estructura del sistema es, como analizaremos posteriormente, siempre deudora de consideraciones pragmáticas.

4. Modelos matemáticos

Una herramienta usual y predominante en las ciencias naturales o las llamadas “ciencias duras” son las ecuaciones matemáticas diseñadas con el fin de describir determinadas propiedades o comportamientos del mundo. Considerando las preeminencia de esta estrategia cognoscitiva se hace imperioso analizar este tipo de modelos para tratar de entender sobre la base de qué los mismos pueden constituirse como representativos.

Resulta intuitivo pensar que la representatividad que podría poseer un modelo matemático no puede ser explicada en términos de similitud, debido a que no puede considerarse que una ecuación, tal como se presenta en su forma escrita, tenga alguna semejanza con el fenómeno que pretende describir. Vale mencionar que no sería correcto interpretar “similar” como “similar forma o estructura”, ya que el criterio de la similitud se instituye como un criterio débil o laxo precisamente a fin de evitar los problemas que surgen al entender la representación bajo la rigidez de los parámetros del isomorfismo.

Aquello que parece justificar la representación obedece a la relación existente entre la estructura matemática del modelo y la estructura física del fenómeno de estudio. El problema surge cuando se recuerda que la ecuación matemática es un objeto formal y no un sistema físico determinado. En este sentido resultaría inapropiado confundir entidades matemáticas con entidades físicas, e inferir, por ejemplo, de las propiedades del círculo propiedades de la Tierra (Fahmi, 2010). Como afirma Frigg (2002, p. 13), los modelos matemáticos involucran una especie de *realización múltiple*, debido a que una misma ecuación matemática puede ser usada en diferentes contextos modelando distintos sistemas.

Ahora bien, una ecuación puede ser isomorfa a diferentes sistemas, ¿pero será modelo de todas ellas? Si es esencial a la representación el ser representación de algo en particular, el representar exigiría una relación de “uno a uno” no compatible con la realización múltiple que suponen los modelos matemáticos. La estructura supuestamente descrita por la ecuación no refiere por sí misma a nada en particular y, por lo tanto, no puede instituirse como condición suficiente de la representación. Las matemáticas, en y por sí mismas, nada nos dicen de la realidad.

Por otro lado, como analizaremos con mayor detenimiento en el siguiente apartado, en el contexto de discusión de los “falsos” modelos, los problemas propios a las idealizaciones inherentes a los modelos físicos tridimensionales tienen su paralelismo en el ámbito de las matemáticas. Las ecuaciones que pretenden describir la dinámica de algún sistema son siempre *aproximaciones*. Como sostiene Suárez (2003, p. 27): “Podría decirse que la mecánica newtoniana no describe ningún movimiento planetario actual de un modo cuantitativamente preciso”. Aun cuando los inconvenientes acarreados por las aproximaciones son menos apremiantes que los derivados de las idealizaciones, ya que una aproximación es una inexactitud y no una distorsión de lo que sucede, no por ello deja de ser intrigante la inadecuación empírica de los modelos matemáticos. Mientras que una ecuación específica cuantitativamente cierta propiedad, el objeto correspondiente del sistema siempre tendrá un valor diferente.

5. Modelos idealizados

Todos los modelos se construyen sobre la base de la abstracción de aquellas propiedades, entidades o relaciones que se consideran irrelevantes a la luz de los

elementos de interés. La simplificación, al mitigar las complejas interacciones del mundo real, facilita la manipulación, el cálculo y el análisis del objeto de estudio. Tal procedimiento es de vital importancia a la hora de considerar las potencialidades cognoscitivas de los modelos (Kennedy, 2010). En este sentido, tematizar, aunque sea brevemente, en torno al problema suscitado por las idealizaciones se vuelve crucial

Quienes fundamentan el valor epistémico de los modelos en la capacidad representativa, pueden solventar con relativo éxito el proceso de *abstracción* debido a que, en tales casos, el elemento abstraído se mantiene indeterminado (Contessa, 2006). Sin embargo, la *idealización*, entendida en términos de una deliberada distorsión al adscribir una propiedad no correspondida en el sistema, genera serios inconvenientes.

La perspectiva representacionista sustentada sobre las bases de la similitud puede darnos algún tipo de explicación de los inexactos, inadecuados o incluso “falsos” modelos que encarnan las idealizaciones, en tanto que lo único que exige es que ambos dominios sólo compartan *algunas* propiedades (Giere, 2011). En cambio, el isomorfismo es incapaz de dar cuenta en sus propios términos de las representaciones “inadecuadas” o *misrepresentations*: su formalismo, al suponer una biyección entre el modelo y el sistema, conduce a que o bien el modelo es una descripción estructuralmente adecuada del fenómeno en cuestión, o bien no es una representación en absoluto. La explicación vía isomorfismo, en contraposición a lo que sucede en la práctica científica, no puede identificar la “falsa” representación como representación. Las versiones debilitadas que, apelando a diferentes tipos de morfismos como el homomorfismo o el isomorfismo parcial, pretenden dar cuenta de tales circunstancias instituyen débiles e inútiles condiciones para la representación. Como afirma Gabriele Contessa (2006, p. 375), cualesquiera dos estructuras parciales son trivialmente isomórficas siempre que los aspectos discordantes se mantengan indeterminados en los respectivos sistemas. En este sentido, el isomorfismo parcial es insuficiente, ya que con él cualquier modelo representa cualquier sistema. Por lo tanto, ante los problemas suscitados por los estrictos criterios impuestos por el isomorfismo, modificar el “morfismo” por algún homomorfismo o isomorfismo parcial no pareciera ser una buena estrategia pues, debilitando el tipo de relación entre los dominios, el estructuralismo contrae los mismo problemas que los que abogaban por un criterio no formal como el de la similitud. En efecto, tal como se desprende de lo que afirma Contessa, así como la similitud, el isomorfismo parcial o el homomorfismo, en tanto triviales, se instituyen como condiciones tanto de la representación como de la no representación.

En el siguiente apartado analizaremos el problema generado ante la existencia de modelos exitosos pero incompatibles. El caso a presentar no será desarrollado en detalle pues este trabajo ya ha sido realizado en “Entre mecánica cuántica y estructuras químicas: ¿a qué refiere la química cuántica?” por Olimpia Lombardi y Juan Camilo Martínez González. Sin embargo creemos que brindaremos los elementos necesarios para comprender en qué sentido la existencia de modelos incompatibles pone en cuestión cualquier postura representacionista.

6. Un problema insoslayable: modelos incompatibles

La existencia de modelos pragmáticamente exitosos que, refiriéndose al mismo sistema, son incompatibles entre sí no sólo anula la inferencia que extiende tesis realistas a partir del éxito predictivo, sino que consecuentemente cuestiona el valor epistémico de los modelos en términos de representación. Como afirma Margaret Morrison (2011) los

casos en consideración no comportan los mismos problemas epistémicos que suscitan los “falsos” modelos. Diferentes estrategias aproximativas o des-idealizantes contribuyen a acortar las distancias entre éstos y el sistema, generando versiones más realistas respecto del fenómeno descrito. Por el contrario, ante modelos incompatibles resulta muy difícil esgrimir salvoconductos que nos permitan determinar cómo es que ambos modelos nos dicen algo de la realidad en términos de representación. Tanto el isomorfismo como la similitud se manifiestan impotentes: en tanto incompatibles, los modelos no pueden ser ni isomorfos ni similares entre sí (al menos en el aspecto relevante que hace a la mentada incompatibilidad) y, consecuentemente, tampoco pueden ser ambos isomorfos o similares respecto del sistema.

Un ejemplo sumamente interesante que puede ser ilustrativo para el caso de modelos incompatibles utilizados ampliamente en la práctica científica es el que se da en el ámbito de la química cuántica, donde se utilizan modelos que incorporan elementos provenientes tanto de la física clásica como de la mecánica cuántica. Esta peculiaridad torna conflictiva la búsqueda de los referentes de la disciplina (Labarca & Lombardi, 2010; Lombardi & Martínez González, 2012).

En química cuántica se formularon dos enfoques irreductibles para dar cuenta de la naturaleza del enlace químico en una molécula. El concepto de enlace químico, que refiere al fenómeno que mantiene unidos y estabilizados los componentes de las moléculas, constituye un concepto fundamental en química estructural debido a que, al determinar la estructura de las moléculas a partir de los enlaces atómicos, organiza el conocimiento acerca de las sustancias. Fue Gilbert Lewis (1916) quien desarrolló la primera teoría del enlace químico en el marco de la regla del octeto. Esta regla señalaba la tendencia de los átomos a completar sus últimos niveles de energía con ocho electrones. Desde esta perspectiva, el enlace químico se definía en términos de los electrones compartidos por dos núcleos atómicos a fin de satisfacer, en la capa de valencia, lo que la regla prescribía (para el desarrollo histórico del concepto de enlace químico, ver Sutcliffe, 1996).

Frente a algunas dificultades de la teoría de Lewis, y con el advenimiento de la mecánica cuántica en la década de 1920, surge la química cuántica intentando explicar los resultados de la química estructural en términos de la nueva teoría (Woody, 2000). Es precisamente en el marco de la química cuántica que se inscriben las “teorías” del *Enlace de Valencia* (EV) y del *Orbital Molecular* (OM). Estos enfoques, cuya connotación *teórica* no tiene más que la norma consuetudinaria del nombrarlos como ‘teorías’, deberían ser interpretados más bien como modelos. Como afirman Olimpia Lombardi y Juan Camilo Martínez González (2012), ni EV ni OM pueden constituirse como teorías autónomas ya que ambas incorporan el mismo constructo teórico, la ecuación de Schrödinger independiente del tiempo, proveniente de la mecánica cuántica. En tanto co-dependientes de la mecánica cuántica, no puede decirse que posean un marco teórico propio e independiente. En esta dirección, ya en 1939, Slater propuso interpretarlos al modo de modelos, como métodos de aproximación que difieren “no en sus fundamentos sino en la naturaleza precisa de los pasos analíticos utilizados” (Slater, 1939, p. 368). Por lo tanto, EV y OM, en tanto que comparten los mismos insumos teóricos, serían modelos que describen de un modo distinto la estructura interna de la molécula.

Lo particularmente relevante, y por lo cual la apelación al caso se vuelve pertinente, es que tales modelos describen de un modo diferente e incompatible la estructura molecular. En efecto, EV y OM, compartiendo el mismo formalismo matemático, se constituyen como dos métodos de búsqueda alternativos para resolver la

ecuación de Schrödinger. Esta ecuación sólo provee a la química del formalismo de la mecánica cuántica, pero sin ninguna instrucción para modelar moléculas desde esta nueva perspectiva (cf. Woody, 2000). En este sentido, los diferentes métodos de búsqueda dependieron de ciertas *decisiones* conceptuales respecto del sistema molecular, decisiones que justificaban el algoritmo particular para cada uno de los modelos. Mientras que para EV la molécula es una entidad compuesta, donde es posible continuar identificando los átomos componentes, para OM la molécula es un todo no susceptible de ser analizado a partir de componentes atómicos, ya que en esta “nueva” entidad ni los átomos pueden ser identificados, ni los electrones se ubican en torno a un núcleo particular. EV permite representar las posiciones de los electrones en las moléculas, describiendo el compuesto molecular como un conjunto de átomos donde los electrones son localizables y se encuentran orbitando átomos particulares. En este sentido, EV no se aparta radicalmente de la química estructural, sino que se considera como una herramienta que, siguiendo la manera estructural de pensar los problemas químicos, utiliza el formalismo de la mecánica cuántica (Pauling & Wilson, 1935; Park, 1999; Hendry, 2006). En cambio, OM, se aleja de la química estructural ya que presupone una especie de holismo molecular, implicando, con ello, una nueva entidad conceptual donde los electrones no están localizados en orbitales atómicos sino en orbitales moleculares deslocalizados alrededor de la molécula entera.

Si bien EV y OM presentan descripciones moleculares diferentes e incluso incompatibles, coexisten al interior de la química cuántica. En efecto, actualmente no existe una instancia inapelable de decisión entre ambos enfoques. En los orígenes de la química cuántica predominó el enfoque EV, debido a que su familiaridad con la química estructural permitía una representación visual que garantizaba un marco de aplicabilidad (Brush, 1999). Actualmente el enfoque OM tiene preeminencia en el ámbito de la química cuántica computacional (Shaik & Hiberty, 2008). No obstante, ello no significa un debilitamiento de EV, puesto que su aplicación continúa vigente en los ámbitos de la fotoquímica y de la química del estado sólido (Shaik & Hiberty, 2008).

La existencia de estos modelos incompatibles, en tanto que plantean modos de concebir la molécula diametralmente diferentes e irreductibles, pone en cuestión no sólo la posibilidad de encontrar una propiedad que legitime la representación, sino, y fundamentalmente, la concepción según la cual la riqueza epistémica de la producción científica descansa en sus capacidades representativas. En efecto, el uso fructífero y el éxito pragmático de ambos modelos, en conjunción con el hecho de que el sistema descrito por los modelos no puede tener al mismo tiempo las características por ellos descritas, obliga a preguntarnos por cómo es el sistema que, en definitiva, la química cuántica, a través de EV y OM, pretende conocer.

Quizás podría decirse que, ante tales situaciones, sería legítimo preguntarse por cuál de las dos perspectivas, de los dos modelos, representa más adecuadamente al sistema. El sentido de tal pregunta encubre la creencia, absolutamente irrefutable en tanto creencia, que uno y sólo uno puede ser el correcto (también, bajo la misma creencia, podría suponerse que los dos son incorrectos). Es claro que, al ser incompatibles entre sí, no pueden ser verdades parciales que se complementan para dar cuenta del fenómeno en términos de una explicación totalizadora que comprenda todos los escorzos. Por lo tanto, el realista se encuentra obligado a tomar la decisión de considerar, a lo sumo, uno de ellos como el modelo objetivo. En este sentido, uno sería un modelo realista, el otro, en el mejor de los casos, un modelo instrumentalista.

El problema de tal interpretación radica en que no hay motivos internos suficientes para legitimar tal decisión. En efecto, lo perturbador de la situación planteada para analizar el estatus epistémico de los modelos, es que ambos casos satisfacen los criterios necesarios que toda concepción representacionista y realista pareciera exigir para la aceptación de los modelos propuestos en ciencia: ambos son empíricamente adecuados; ambos establecen predicciones exitosas; y ambos permiten realizar inferencias sobre el sistema. Es en este sentido que las motivaciones para preferir uno de los modelos por sobre el otro no están legitimadas por los criterios de corrección estándar que la ciencia utiliza para la aceptación de una teoría, sino que emergen como consecuencia de ciertos supuestos no justificados.

7. El giro pragmático

En los apartados precedentes se ha intentado poner de manifiesto las limitaciones de los intentos de dar cuenta de la representación en términos de una propiedad compartida entre los modelos y el mundo. El derrotero a través de los diferentes tipos de modelos expuso las dificultades de toda versión que naturalice la representación e intente explicarla exclusivamente a partir de una relación diádica entre ambos dominios. La pretensión de fundamentar el vehículo de la representación en el isomorfismo o la similitud no logra abarcar todos los tipos de modelos, ni establecer las condiciones necesarias y suficientes del representar. En efecto, ambas versiones, revelando su incapacidad para establecerse como fuentes en y por sí mismas del representar, carecen de la direccionalidad que supone el ser representación de algo en particular. Por un lado, en tanto que cualesquiera dos cosas son similares en algún aspecto, bajo este criterio o bien todo representa a todo, o nada representa nada. Por otro lado, el isomorfismo, en primer lugar, supone incorrectamente que o bien el mundo ya se encuentra *per se* estructurado, o bien puede estructurárselo de un modo definido. En segundo lugar, el isomorfismo tampoco contiene una direccionalidad de modo inherente a la estructura ya que ésta, como dijimos precedentemente, puede ejemplificarse en diferentes sistemas.

Ante tales dificultades, Giere propone una concepción pragmatista respecto de los modelos científicos pensando la lógica de la representación a partir de cuatro elementos: un *agente* que usa un *modelo* para representar un sistema para algún *propósito* determinado (Giere, 2004; 2010). Desnaturalizando la fuente de la representación, el pragmatismo adopta una relación triádica al sumar al binomio mundo-modelo al agente que, constituyéndose como elemento vinculante, garantiza la direccionalidad del modelo hacia el sistema a través de su propia intencionalidad. De este modo se desvanecería la crítica respecto de la insuficiencia de la similaridad o del isomorfismo para vincular modelo y mundo ya que, precisamente, sería el científico quien, determinando los caracteres relevantes para la representación, establecería tal conexión (Giere, 2011).

Ahora bien, bajo la concepción pragmatista la potencialidad representativa del modelo radicaría, en desmedro de cualquier rasgo intrínseco de éste, en la actividad desarrollada por el agente. En este sentido, parece que la noción de representación basada en la intencionalidad es demasiado liberal como para caracterizar a la representación científica. Por un lado, un hablante competente y coherente no puede pretender adscribir potencialidad representativa a dos modelos incompatibles. Por otro lado, si bien la versión pragmatista tiene el mérito de incorporar a la discusión al sujeto de la acción del representar, a partir de ella no es muy claro qué puede decirse respecto de la representación. Si consideramos que la utilidad de la pregunta por la representación

estriba en que es una de las formas de explicar la capacidad epistémica de los modelos, el pragmatismo no puede decir demasiado respecto a que los modelos brindan conocimiento *porque* representan al sistema. El conocimiento obtenido no sería producto de que el modelo *represente* al sistema, sino de que el sujeto *utilice* al modelo con fines representativos. Reduciendo el poder representativo a la intencionalidad, el pragmatismo no nos informa respecto a los fundamentos que justifican la representación, ni cómo es que los modelos nos brindan conocimiento vía representación. Como afirma Frigg (2002, p. 19):

Cuando nos preguntamos cómo funciona la representación, lo que nos gustaría saber es qué hace exactamente el científico cuando usa S para representar T. Si decimos que intenta representar T por medio de S, meramente parafraseamos el problema y no lo respondemos, porque lo que queremos saber es qué involucra dicha intención.

Pero entonces, ¿qué involucra dicha intención? En su trabajo *Scientific Representation: Against Similarity and Isomorphism*, Suárez propone una concepción inferencial de la representación, según la cual A representa B sólo si (i) la fuerza representacional de A nos orienta o apunta a B (o sea, si A es capaz de guiar a un usuario competente e informado para poner en consideración al sistema B), y (ii) A permite hacer inferencias científicas sobre B. La propuesta de Suárez es una concepción deflacionaria, diseñada en vistas a desentenderse de los detalles en los debates entre los partidarios del isomorfismo y la similitud. Con ello se intenta evitar la identificación de alguna propiedad en particular que caracterice la representación. Para tal fin, meramente revaloriza la capacidad de establecer inferencias desde el modelo al sistema.

En concordancia con Suárez, podríamos afirmar que la potencialidad epistémica de los modelos reside en su capacidad de establecer inferencias respecto del sistema. Ahora bien, sobre la base de las dificultades expuestas en las secciones anteriores, queda de manifiesto que la propiedad inferencial de los modelos no puede ser entendida como una propiedad intrínseca y natural de los mismos, sino que depende de ciertos criterios pragmáticos de selección de variables y de supuestos que condicionan y legitiman el proceso de modelado. Por ejemplo, cuando analizamos el caso de EV y OM descubrimos que cada uno de estos métodos contiene un algoritmo particular para generar funciones de onda, y la justificación para elegir ese algoritmo descansa en supuestos conceptuales y cualitativos acerca de la forma que debería tener la función de onda molecular. Cuando Walter Heitler y Fritz London (1927) formularon el enfoque EV al intentar resolver la ecuación de Schrödinger independiente del tiempo, partieron de una función de onda electrónica construida para representar la molécula de hidrógeno H_2 . Pero para encontrar dicha función de onda tuvieron que suponer al sistema como compuesto de dos núcleos y dos electrones completamente separados, y sólo sobre esta base pudieron construir la función de onda correspondiente. A su vez, también el éxito del enfoque OM descansa sobre dos supuestos fundamentales. El primer supuesto, llamado campo autoconsistente (SCF por sus siglas en inglés), considera que, para un electrón cualquiera, su distribución de carga media es exactamente igual a la que se obtiene cuando se resuelve la ecuación de Schrödinger con un campo potencial debido a las cargas del núcleo y al promedio de las cargas de los otros electrones. El segundo supuesto afirma que cuando un electrón está cerca de un núcleo particular en un sistema molecular, todas las fuerzas que actúan sobre él son debidas a ese núcleo y a los electrones cercanos al

mismo. Como fue señalado, a pesar de la profunda diferencia en las imágenes que los dos modelos brindan de la molécula, ambos se utilizan de un modo exitoso en la práctica de la química cuántica.

La peculiaridad del caso analizado de modelos incompatibles pone de manifiesto que, si bien la capacidad inferencial puede constituirse como requisito necesario para la potencialidad representativa, no constituye una condición suficiente para ella. En efecto, en tanto que ambos modelos adscriben propiedades incompatibles al mismo sistema, resulta difícil atribuir capacidad representativa a los modelos a partir de su éxito pragmático: puesto que ambos modelos son incompatibles entre sí, aun estableciendo las inferencias exitosas, no podemos decir cómo es el sistema. Por lo tanto, parece más acertado adoptar una posición más prudente que la de Suárez, afirmando que en algunos contextos la capacidad inferencial juega un papel representativo y en otros casos cumple un papel eminentemente instrumental. En este sentido la química cuántica nos brinda la posibilidad de pensar de un modo distinto la forma en que la ciencia produce conocimiento. En efecto, en tal disciplina el conocimiento se genera a partir de la capacidad de cálculo y predicción de los modelos; cálculos y predicciones no siempre susceptibles de ser comprendidos desde marcos epistemológicos representacionistas.

8. Conclusiones

En el presente trabajo se intentó pensar de qué modo ciertas versiones tradicionales de la representación, como lo son la postura estructuralista pensada desde un isomorfismo o isomorfismo parcial y la similitud, podrían explicar el proceso representativo considerando la multiplicidad de los modelos existentes en la práctica científica. Como se pretendió poner de manifiesto, desde una perspectiva naturalista del representar, se presentan serias dificultades filosóficas cuando intentamos justificar su pertinencia pensándola a partir de aquellos modelos en donde apretantemente se aplicarían. Asimismo ninguna de ellas puede afrontar exitosamente todas las circunstancias específicas que plantean las particularidades de cada tipo de modelo. A su vez, el giro pragmatista, si bien constituye un avance en la medida en que desnaturaliza la relación de representación al considerar el papel que cumple el agente, se enfrenta a serios problemas ya sea, en primer lugar, por deflacionar el concepto de representación limitándolo a la capacidad inferencial del modelo, o, en segundo lugar, cuando se trata de explicar la representatividad de modelos exitosos pero incompatibles. En relación a estas consideraciones hemos evaluado la posibilidad de reducir el alcance de la propuesta de Suárez afirmando que la capacidad inferencial no puede constituirse en una condición suficiente del representar ya que en algunas situaciones la capacidad inferencial de algunos modelos cumple solamente funciones instrumentales.

Como conclusión parcial podría afirmarse que prescindir del agente intentando estipular un criterio plenamente naturalista de la representación a partir exclusivamente de las propiedades del modelo parece tan imprudente como pretender captar un único criterio para la representación ante semejante diversidad de tipos de modelos. Es necesario estudiar cada modelo en particular para determinar en cada caso su potencialidad representativa y en virtud a qué propiedad tal representación se instituye. Los criterios generales acerca de la representatividad de los modelos siempre encontrarán limitaciones a la luz de las múltiples y disímiles maneras en que los modelos son utilizados de manera fecunda en ciencias.

9. Agradecimientos

Agradecemos al Grupo de Filosofía de las Ciencias dirigido por Olimpia Lombardi en la Universidad de Buenos Aires al cual pertenecemos, por su constante respaldo. Este trabajo ha sido posible gracias al apoyo económico del subsidio N.º 57919 otorgado por la John Templeton Foundation.

10. Referencias

- Brush, S. (1999). Dynamics of theory change in chemistry: Part 2. Benzene and molecular orbitals, 1945-1980. *Studies in History and Philosophy of Science*, 30 (2), 263-302.
- Contessa, G. (2006). Scientific models, partial structures and the new received view of theories. *Studies in History and Philosophy of Science*, 37 (2), 370-77.
- Fahmi, M. (2007). Mathematical models in science: a debate about ontology. *Conference Models and Simulations 2*, Tilburg 11.10.2007 - 13.10.2007. Recuperado de <http://philsci-archive.pitt.edu/3904/>.
- Frigg, R. (2002). Models and representation: why structures are not enough. En P. Dietsch (Ed.), *Measurement in Physics and Economics. Discussion Papers*. London, UK: The London School of Economics and Political Science.
- Frigg, R. (2006). Scientific representation and the semantic view of theories. *Theoria*, 21 (1), 49-65.
- Frigg, R. & Hartmann, S. (2006). Models in science. En E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Stanford, USA: Stanford University. Recuperado de <http://plato.stanford.edu/entries/models-science/>.
- Giere, R. (2004). How models are used to represent physical reality. *Philosophy of Science*, 71 (5), 742-52.
- Giere, R. (2010). An agent-based conception of models and scientific representation. *Synthese*, 172 (2), 269-81.
- Giere, R. (2011). Representing with physical models. En P. Humphreys y C. Imbert (Eds.), *Models, Simulations and Representations*. New York, USA: Routledge.
- Harris, T. (2003). Data Models and the Acquisition and Manipulation of Data, *Philosophy of Science*, 70 (5), 1508-17
- Heitler, W. & London, F. (1927). Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. *Zeitschrift für Physik*, 44, 455-72.
- Hendry, R. F. (2006). Two conceptions of the chemical bond. *Philosophy of Science*, 75 (5), 909-20.
- Kennedy, A. G. (2010). Models and Scientific Explanation. *Philosophy of Science Association. 22nd Biennial Meeting, Montreal*. Recuperado de <http://philsci-archive.pitt.edu/id/eprint/8374>.
- Knuuttila, T. (2005). Models, representation and mediation. *Philosophy of Science*, 72 (5), 1260-71.

- Knuuttila, T. (2009). Representation, idealization, and fiction in economics. From the assumptions issue to the epistemology of modeling. En M. Suárez (Ed.), *Fictions in Science: Philosophical Essays on Modeling and Idealization*. London, UK: Routledge.
- Labarca, M. & Lombardi, O. (2010). Acerca del status ontológico de las entidades químicas: el caso de los orbitales atómicos. *Principia*, 14 (3), 309-33.
- Lewis, G. N. (1916). The atom and the molecule. *Journal of the American Chemical Society*, 38 (4), 762-85.
- Lombardi, O. & Martínez González, J. C. (2012). Entre mecánica cuántica y estructuras químicas: ¿a qué refiere la química cuántica?. *Scientiae Studia*, 10 (4), 649-70.
- Morrison, M. (2011). One phenomenon, many models: Inconsistency and complementarity. *Studies in History and Philosophy of Science*, 42 (2), 342-51.
- Park, B. S. (1999). Chemical translators: Pauling, Wheland and their strategies for teaching the theory of resonance. *The British Journal for the History of Science*, 32 (1), 21-46.
- Pauling, L. & Wilson, E. B. (1935). *Introduction to Quantum Mechanics with Applications to Chemistry*. New York, USA: McGraw-Hill.
- Shaik, S. & Hiberty, P. (2008). *A Chemist's Guide to Valence Bond Theory*. New York, USA: Wiley.
- Slater, J. C. (1939). *Introduction to Chemical Physics*. London, UK: McGraw-Hill.
- Suárez, M. (2003). Scientific representation: Against similarity and isomorphism. *International Studies in the Philosophy of Science*, 17 (3), 225-244.
- Suppes, P. (1962). Models of data. En E. Nagel, P. Suppes y A. Tarsky (Eds.), *Logic, Methodology and Philosophy of Science: Proceeding of the 1960 International Congress*. Stanford, USA: Stanford University Press.
- Sutcliffe, B. T. (1996). The development of the idea of a chemical bond. *International Journal of Quantum Chemistry*, 58 (6), 645-55.
- Woody, A. I. (2000). Putting quantum mechanics to work in chemistry: the power of diagrammatic representation. *Philosophy of Science*, 67 (3), S612-27.

La cuestión de la Epistemología Histórica como estilo epistemológico

Marcela Becerra Batán¹

Recibido: 28 de marzo de 2016.

Aceptado: 2 de mayo de 2016.

Resumen. En este trabajo planteamos si la Epistemología Histórica –*Épistémologie Historique* (EH) e *Historical Epistemology* (HE)– puede ser entendida como “estilo epistemológico” y nos interrogamos acerca de las principales implicancias de esta cuestión en debates epistemológicos actuales. En tal sentido, primeramente intentamos reconstruir el programa de la EH/HE en sus rasgos distintivos y en su periodización, así como evaluar su fecundidad. Procuramos luego esclarecer la noción de “estilo”, siguiendo los desarrollos de Gayon (1996), de Braunstein (2002, 2006 y 2008) y especialmente de Hacking (1982, 1990, 1992, 2002a, 2002b, 2010 y 2012). Finalmente abordamos la cuestión de la EH/HE como estilo epistemológico, planteamos relaciones entre estilos en epistemología y los “estilos de pensamiento & acción científicos” (Hacking, 2010) y destacamos algunas implicancias de esta cuestión para la práctica actual de la epistemología, particularmente para aquellos ejercicios epistemológicos que buscan situarse “entre un análisis conceptual y una historia de las prácticas científicas” (Vagelli, 2014).

Palabras claves: epistemología histórica – estilos

Title: The issue of Historical Epistemology as epistemological style

Abstract. In this paper we propose whether Historical Epistemology –*Épistémologie Historique* (EH) and *Historical Epistemology* (HE)– can be understood as “epistemological style” and we wonder about the major implications of this issue in current epistemological debates. In this regard, first we try to reconstruct the program EH / HE in its distinctive features and its periodization, and to assess their fertility. Then we try to clarify the notion of “style” by following the contributions of Gayon (1996), Braunstein (2002, 2006 and 2008), and especially Hacking (1982, 1990, 1992, 2002a, 2002b, 2010 and 2012). Finally we address the issue of EH / HE as epistemological style, propose relationships between styles in epistemology and “styles of scientific thinking & doing” (Hacking, 2010) and highlight some implications of this issue for the current practice of epistemology, particularly for those epistemological exercises that seek to be “between conceptual analysis and history of scientific practices” (Vagelli, 2014).

Keywords: historical epistemology – styles.

1. Introducción

En este trabajo planteamos si la Epistemología Histórica –*Épistémologie Historique* (EH) e *Historical Epistemology* (HE)– puede ser entendida como “estilo epistemológico” y

¹ Universidad Nacional de San Luis, Facultad de Ciencias Humanas.

✉ mbatan2@gmail.com

Becerra Batán, Marcela; (2016); “La cuestión de la Epistemología Histórica como estilo epistemológico”; *Epistemología e Historia de la Ciencia*; 1(1); 35-52. ISSN: 2525-1198



nos interrogamos acerca de las principales implicancias de esta cuestión en debates epistemológicos actuales. En tal sentido, primeramente intentamos reconstruir el programa de la EH/HE en sus rasgos distintivos y en su periodización, así como evaluar su fecundidad. Procuramos luego esclarecer la noción de “estilo”, siguiendo los desarrollos de Gayon (1996), de Braunstein (2002, 2006 y 2008) y especialmente de Hacking (1982, 1990, 1992, 2002a, 2002b, 2010 y 2012). Finalmente abordamos la cuestión de la EH/HE como estilo epistemológico, planteamos relaciones entre estilos en epistemología y los “estilos de pensamiento & acción científicos” (Hacking, 2010) y destacamos algunas implicancias de esta cuestión para la práctica actual de la epistemología, particularmente para aquellos ejercicios epistemológicos que buscan situarse “entre un análisis conceptual y una historia de las prácticas científicas” (Vagelli, 2014).

2. La cuestión de la Epistemología Histórica

La EH/HE constituye actualmente un campo en crecimiento, con líneas de investigación divergentes. Surgida a partir de Comte, la EH tuvo un período singularmente innovador y potente en el ámbito filosófico francés con Bachelard, Canguilhem y Foucault. A partir de los años noventa, esta perspectiva viene siendo retomada y resignificada por filósofos como Hacking, Daston, Rheinberger y Davidson entre otros, cultivada más allá de Francia en instituciones de otros países de Europa, de América del Norte y de América del Sur y notablemente vuelta a valorizar hoy en su propia tierra natal (cf. AAVV, 2012 & Vagelli, 2014). Cabe entonces primeramente reconstruir este programa en sus rasgos distintivos y en su periodización, así como evaluar su fecundidad.

De la compleja historia de la palabra², rescatamos lo que resume Braunstein (2012): la EH es una vieja expresión para un viejo estilo de pensamiento, el “estilo francés en filosofía de la ciencia”. Estilo que comienza con Comte, continúa con Rey y prosigue con Bachelard, Canguilhem, Althusser y Foucault, y cuyos rasgos distintivos son los siguientes (cf. Braunstein, 2002 & 2012):

- es una epistemología contemporánea de las ciencias, que no intenta determinar un fundamento a priori de las mismas;
- es una epistemología que no es una “teoría del conocimiento” en general y en tal sentido, es una epistemología que critica las ideas de un método único y estático, de un sujeto a-histórico fundador del conocimiento y de un objeto de conocimiento dado e inmediatamente “objetivo”;

² Braunstein (2012) halla por primera vez el término “epistemología histórica” en la tesis doctoral de Abel Rey en 1907, así como en otros textos posteriores del mismo autor. Más tarde, en 1963, Canguilhem usa este término (cf. Canguilhem, 1968). Según lo que Lecourt ha relatado a Gayon (2006), Canguilhem trazaba una distinción entre “epistemología histórica” (con la que designaba el trabajo de Bachelard) e “historia epistemológica” (con la que designaba en cambio su propio trabajo). En 1969, Lecourt escribe su memoria de Maestría titulada La epistemología histórica de Gaston Bachelard (cf. Bibliografía); texto que, junto a otros del mismo autor, contribuiría a la difusión de Bachelard y de la epistemología francesa fuera de Francia. Por otra parte, se han localizado usos de la denominación “EH” en textos marxistas; en tal sentido, Lecourt (2007) halla la expresión en un texto de Marx Wartofsky de 1973, con la cual éste se refería a su metodología marxista leninista en historia de la ciencia. Gingras (2010) encuentra otro texto marxista escrito en 1980 por el lógico polaco Jery Kmita, precisamente titulado: Problemas de Epistemología Histórica.

- es una reflexión histórica sobre las ciencias. La relación entre filosofía y ciencias resulta fructífera cuando está basada sobre una historia de las ciencias, concebida ésta no como una “historia tradicional”, sino como una historia filosófica, crítica, valorizada, juzgada desde la perspectiva del presente, recurrente, regional, discontinua, atenta a los errores, a las ideologías científicas y/o a las relaciones saber-poder;
- es una epistemología para la cual la racionalidad tiene una historia y una geografía, un pasado y una actualidad, en un dominio y en unas determinadas condiciones de ejercicio;
- es una epistemología que no conduce a un relativismo, sino que da cuenta de las condiciones históricas bajo las cuales se puede enunciar algo como verdadero en un discurso científico;
- es una epistemología que se vincula con una historia del presente y con objetivos políticos.

En cuanto a la periodización de la EH / HE, se distinguen en ella dos períodos o etapas (cf. Braunstein, 2012; Méthot, 2012; Vagelli, 2012a y 2012b, p. 1) la “vieja” EH, o la etapa originaria en suelo francés y 2) la “nueva” HE, la etapa contemporánea a partir de los años 90 más allá de Francia, especialmente en el ámbito anglo-americano. Entre ambas etapas, se han señalado continuidades y rupturas.

La etapa nueva o contemporánea se inicia en 1993, a partir de una conferencia internacional dedicada a la HE en la Universidad de Toronto, organizada por Hacking y Daston. Otras conferencias posteriores en Berlín, Columbia, Lovaina, París, así como una serie de publicaciones diversas, se dedicarían al mismo tema se dedicarían al mismo tema.

En esta etapa contemporánea, se destacan las producciones surgidas desde el Instituto Max Planck para la Historia de las Ciencias (MPIWG) de Berlín, que pueden ordenarse según la tipología propuesta por Feest y Sturm (2011):

1. Historias de conceptos epistémicos de alto nivel, como los de objetividad, observación, probabilidad, etc. (ej.: los estudios de Daston);
2. Historias de objetos epistémicos, de las trayectorias históricas de objetos científicos particulares, como el electrón, el ADN, el flogisto, etc. (ej. los estudios de Rheinberger);
3. Dinámicas de desarrollos científicos a nivel de la larga duración, como las transformaciones de los sistemas de conocimiento físico (ej. los estudios de Renn).

Asimismo en esta etapa contemporánea, cabe destacar el espacio de investigación, formación e intercambios constituido en la Universidad Paris 1 *Panthéon-Sorbonne*. Entre otras actividades de interés, desde esta institución se han organizado dos Jornadas de Estudio dedicadas a la EH: las primeras “Jornadas de estudios Epistemología histórica: comienzos y cuestiones actuales” (*Journées d’études Épistémologie historique: commencements et enjeux actuels*) realizadas en mayo de 2015 y las “Segundas Jornadas de estudios Epistemología histórica; una historia del presente” (*Deuxièmes Journées d’études Épistémologie historique; une histoire du présent*), realizadas en mayo de 2016. En este espacio, la EH es entendida “en sentido lato, en tanto que tradición y método en filosofía e historia de las ciencias” (Vagelli, 2015). En esta orientación, se promueven

investigaciones acerca de la historia de la EH (desde su fase originaria hasta su fase contemporánea), debates metodológicos en filosofía e historia de las ciencias en relación con prácticas en diversos dominios disciplinares y en todo ello, reflexiones acerca del sentido crítico del análisis propio de la EH.

Rescatamos a continuación algunas citas claves de representantes actuales que muestran diversas definiciones de la EH/ HE:

La Epistemología histórica apunta a mostrar cómo esta nueva forma de experiencia que llamamos 'sexualidad' está vinculada con la emergencia de nuevas estructuras de conocimiento, y especialmente con un nuevo estilo de razonamiento y los conceptos empleados en él" (Davidson, 2001: XIII). "Quiero concentrarme en la relación entre formas de experiencia y sistemas de conocimiento, en el modo en el cual lo que hemos llegado a llamar 'sexualidad' es el producto de un sistema de conocimiento psiquiátrico que tiene su propio y particular estilo de razonamiento y de argumentación (...) Como Foucault, estoy interesado en cómo sistemas de conocimiento nos forman como sujetos, cómo esos sistemas literalmente nos hacen sujetos (Davidson, 2001, p. 32)³.

Lo que entiendo por *epistemología histórica* es la historia de las categorías que estructuran nuestro pensamiento, que modelan nuestra concepción de la argumentación y de la prueba, que organizan nuestras prácticas, que certifican nuestras formas de explicación y que dotan a cada una de esas actividades de una significación simbólica y de un valor afectivo. Esta epistemología histórica puede (y de hecho, ella lo debe) hacer referencia a la historia de las ideas y de las prácticas, tanto como a la historia de las significaciones y de los valores que constituyen las economías morales de las ciencias (Daston, 2008, p. 367)⁴.

Rheinberger explica que no usa el término "epistemología" como sinónimo de teoría del conocimiento, sino que más bien

El concepto es usado aquí, siguiendo la práctica francesa, para reflexionar sobre las condiciones *bajo* las cuales, y los significados *con* los cuales, las cosas se convierten en objetos de conocimiento. Se centra así en el proceso de generación de conocimiento científico y en los modos en los cuales éste se inicia y se mantiene (...). Este desplazamiento en la constelación del problema

³ "Historical epistemology attempts to show how this new form of experience that we call 'sexuality' is linked to the emergence of new structures of knowledge, and specially to a new style of reasoning and the concepts employed within it" (Davidson, 2001: XIII). "I want to concentrate on the relation between forms of experience and systems of knowledge, on the way in which what we have come to call 'sexuality' is the product of a system of psychiatric knowledge that has its own very particular style of reasoning and argumentation (...) Like Foucault I am concerned with how systems of knowledge shape us as subjects, how these systems literally make us subjects" (Davidson, 2001, p. 32).

⁴ «Ce que j'entends par *historical epistemology* est l'histoire des catégories qui structurent notre pensée, qui modèlent notre conception de l'argumentation et de la preuve, qui organisent nos pratiques, qui certifient nos formes d'explication et qui dotent chacune de ses activités d'une signification symbolique et d'une valeur affective. Cette épistémologie historique peut (et en fait elle le doit) renvoyer à l'histoire des idées et des pratiques, tout autant qu'à l'histoire des significations et des valeurs qui constituent les économies morales des sciences». (Daston, 2008, p. 367).

está al mismo tiempo tanto en el núcleo de la epistemología como en el punto de partida para su historicización (Rheinberger, 2010, pp. 2-3)⁵.

Reconociendo el intento de Rheinberger de “historicizar” la epistemología, Braunstein (2012) expresa:

En efecto me parece que lo menos que una epistemología que se jacta de ser histórica puede hacer es investigar las condiciones históricas de sus propias producciones. No vamos a contentarnos con preguntar qué es la epistemología histórica o cuál es su valor; nos preguntaremos además por lo que ha sido su historia (Braunstein, 2012, p. 33)⁶.

A pesar de algunas críticas, el programa de la EH/ HE no ha dejado de dar muestras de su fecundidad. Una de las críticas más altisonantes ha sido la de Gingras (2010), quien ha afirmado que la denominación “HE” en el campo anglosajón de la historia y la filosofía de la ciencia, es una etiqueta que “ nombra sin necesidad” o más bien que, en todo caso, su necesidad radica en justificar la existencia y los subsidios de instituciones que buscan ocupar un lugar en el competitivo “mercado de ideas” (aquí se refiere especialmente al MPIWG). Se trataría entonces del uso estratégico de una “marca” para dar cohesión, identidad y visibilidad internacional a una institución, y así promover la “circulación circular” de conceptos producidos en sus programas de investigación. Esta etiqueta ha llegado a tener numerosos y diversos significados, resulta inadecuada respecto del problema que pretende nombrar y del método que se propone utilizar y, más allá de sus grandiosas pretensiones, nombra algo que de distintas maneras ya se viene haciendo desde hace largo tiempo. Ante esta crítica, acordamos con Méthot (2012), quien considera que, más allá de las expresiones polémicas antes referidas, los trabajos producidos desde el MPIWG nos han provisto de herramientas conceptuales útiles para el ejercicio epistemológico –Méthot destaca especialmente las de “*epistemic things*” y “*experimental systems*” propuestas por Rheinberger–. Asimismo acordamos con Braunstein (2012): a diferencia de lo expresado por Gingras, la EH/HE no es una marca que nombra sin necesidad, pues “es imposible crear un mercado de la nada, nada puede ser vendido si no responde a cierta necesidad”⁷ (Braunstein, 2012, p. 33). La EH/HE es más bien una marca que funciona pues viene a dar respuesta a una demanda, y es una marca que persiste desde sus comienzos hasta hoy. En efecto, un número creciente de investigadores de diversas latitudes dan pruebas de la vigencia y de la potencia del programa de la EH/HE.

Por nuestra parte, desde nuestra propia situación, hacemos hincapié en la necesidad de un análisis, crítica y reflexión de la EH/HE, que nos permita volver a transitar el camino que va desde su fase originaria hasta su fase contemporánea –

⁵“The concept is used here, following the French practice, for reflecting on the historical conditions *under* which, and the means *with* which, things are made into objects of knowledge. It focuses thus on the process of generating scientific knowledge and the ways in which it is initiated and maintained (...). This shift in the problem constellation is at the same time both at the core of epistemology and the point of departure for his historicization” (Rheinberger, 2010, p. 2-3).

⁶“It indeed seems to me that the least that an epistemology which boasts to be historical can do is to look into the historical conditions of its own productions. We will not content ourselves with asking what historical epistemology is or what it is worth; we will also wonder what its history has been” (Braunstein, 2012, p. 33).

⁷“It is impossible to create a market from nothing, nothing can be sold if it does not answer a certain need” (Braunstein, 2012, p. 33).

deteniéndonos hoy especialmente en Hacking-, para plantear nuevamente las relaciones entre filosofía /epistemología, historia de las ciencias y ciencias, en favor de una epistemología histórica fecunda, vinculada con nuestra actualidad.

3. La cuestión de los estilos

En segundo lugar, intentamos responder la pregunta acerca de qué se entiende por “estilo/s” y tratamos de esclarecer esta noción, siguiendo los desarrollos de filósofos pertenecientes o cercanos a la EH/HE tales como Gayon (1996), Braunstein (2002, 2006 y 2008) y particularmente Hacking (1982, 1990, 1992, 2002a, 2002b y 2012), en lo que este último ha denominado el “proyecto estilos”, en sus distintas formulaciones hasta la actualidad.

Gayon (1996) comienza presentando una cartografía de algunos usos de la palabra “estilo”, para ir desde allí hacia una posible definición. En tal sentido, en un primer momento, localiza dos diferentes usos de la noción de estilo en historiadores de las ciencias: en historias locales (historias de escuelas o grupos de investigación e historias de naciones) y en historias generales. Mientras que las primeras subrayan el aspecto “individualizante”, las segundas muestran el aspecto “universalizante” de los estilos científicos. Gayon se detiene luego en la cuestión de la categoría de estilo en filosofía de las ciencias. Al interrogarse por el significado de “estilo” en los estilos científicos, destaca que en este término se encuentran aquellos dos aspectos antes mencionados o funciones antinómicas ya presentes en la categoría de estilo de las bellas artes. Y al preguntarse por el interés de la noción de estilos científicos para la filosofía de la ciencia, señala que Hacking ha visto en los estilos de Crombie “una solución de reemplazo” al *a priori* kantiano en el problema de los fundamentos de la objetividad. Además, afirma que los estilos de Crombie son “candidatos” al estatus de “*a priori* histórico” en el sentido definido por Foucault. Este breve trabajo anticipa cuestiones que serían luego desarrolladas con mayor complejidad.

Braunstein ha retomado el concepto de estilo ya elaborado por Fleck, Crombie, Gayon y Hacking, ha realizado diversos análisis del mismo y ha extendido su uso para poder dar cuenta de un “estilo francés” en filosofía y en historia de las ciencias. En este sentido, en Braunstein (2002), considera adecuado describir los caracteres comunes a Bachelard, Canguilhem, Althusser y Foucault no como “escuela” o “tradicción”, sino más bien “en términos de ‘estilo de pensamiento científico’, en el sentido que filósofos de las ciencias como Ludwig Fleck, luego Alistair Crombie o Ian Hacking han podido dar a ese término” (Braunstein, 2002, p. 3)⁸. Retomando a Gayon (1996), para quien el estilo tiene a la vez una función universalizante y otra individualizante, Braunstein se sirve de la noción de estilo para definir “rasgos comunes que aparecen en un momento determinado y perduran un cierto tiempo” –los hemos sintetizado más arriba–, como asimismo para individualizar a autores que ilustran un estilo.

En Braunstein (2006), realiza un análisis comparado de algunas nociones y conceptos claves de Fleck, Canguilhem y Foucault. Particularmente nos interesa la relación que entabla entre los “estilos de pensamiento” de Fleck y los “*a priori* históricos” de Foucault. En este sentido señala que Fleck, anticipando de alguna manera a Foucault, ya concebía a los estilos de pensamiento como fuerzas que constriñen al pensamiento,

⁸ “En termes de ‘style de pensée scientifique’, au sens que des philosophes des sciences comme Ludwig Fleck, puis Alistair Crombie ou Ian Hacking ont pu donner à ce terme” (Braunstein, 2002, p. 3).

como disposiciones para ver, aprehender y percibir de una determinada manera, y no de otra, en una época. Los diferentes modos de pensar y de ver de científicos de distintas épocas dependen, pues, de los diferentes estilos de pensamiento en los que están situados. De manera semejante, los *a priori* históricos de Foucault delimitan campos de saber y definen lo que allí se puede ver, enunciar y pensar como objeto de conocimiento en una época dada. Braunstein selecciona un pasaje (Foucault, 1969, p. 47), en el que Foucault aproxima la noción de *a priori* histórico a la de estilo: al brindar el ejemplo de la ciencia médica del siglo XIX, expresa que ésta se caracterizaba por un *estilo*, un “carácter constante” de la enunciación, una misma mirada, un mismo análisis y un mismo sistema de transcripción de lo perceptible en lo enunciable.

En Braunstein (2008), hace referencia a un “estilo francés” en historia de las ciencias, que va de Comte a Foucault, pasando por Bachelard y Canguilhem (cf. Braunstein, 2008, p. 12). Nuevamente aquí se sirve de la noción de estilo de pensamiento científico de Fleck, Crombie, Gayon y Hacking para poner de manifiesto tanto similitudes y rasgos comunes de un estilo, así como diferencias y rasgos individuales dentro del mismo. Los rasgos comunes del estilo francés en historia de las ciencias, necesariamente vinculados con los rasgos comunes de este estilo en epistemología, son los siguientes:

- la historia de las ciencias está inseparablemente entrelazada con la filosofía de las ciencias;
- la historia de las ciencias es una historia filosófica o crítica, juzgada, recurrente, en la que se abordan las relaciones entre ciencias e ideologías y/o entre saberes y poderes;
- la historia de las ciencias es el punto de partida para una reflexión sobre la historicidad de la racionalidad.

Podría agregarse que este estilo francés en historia de las ciencias plantea relaciones entre ciencia y sociedad, pero sin caer en una “sociologización” de las ciencias. En una entrevista reciente, refiriéndose a las relaciones entre ciencia y sociedad en el enfoque de Canguilhem y en el de otros autores de este estilo, Braunstein expresa:

Hay evidentemente un peso de la sociedad o de la cultura sobre la ciencia. Las ciencias están en su época, pero al mismo tiempo no se reducen a esto. Desde este punto de vista, hay por lo tanto una crítica de los *Science and technology studies* o de la idea de una ‘sociologización’ de las ciencias. Las ciencias están ligadas a la sociedad, pero sin embargo no son el puro resultado de la sociedad (Almeida & Camelozi, 2016, p. 174)⁹

El trabajo de Braunstein en torno al concepto de estilo en epistemología e historia de las ciencias brinda claves y orientaciones que han sido y son fructíferas para promover indagaciones en el campo actual de la EH/HE.

Hacking, por su parte, ha elaborado, rectificado y aplicado su concepto de estilos desde fines de los ´70 hasta la actualidad. En Hacking (1982) presenta por primera vez su versión de los “estilos de razonamiento científico”, que tomó, para sus propósitos

⁹ « Il y a évidemment un poids de la société ou de la culture sur la science. Les sciences sont dans leur époque, mais en même temps elles ne se réduisent pas à cela. De ce point de vue-là, il y a donc une critique des *Science and technology studies* ou de l'idée d'une 'sociologisation' des sciences. Les sciences sont liées à la société, mais néanmoins elles ne sont pas le pur résultat de la société » (Almeida y Camelozi, 2016, p. 174).

filosóficos, de los “estilos de pensamiento científico” o distintos “métodos de investigación en la tradición europea occidental” propuestos por Crombie en (1978, p. 1)- la simple postulación en ciencias matemáticas, 2- la exploración experimental y medición de relaciones observables, 3- la construcción hipotética de modelos analógicos, 4- el ordenamiento de la variedad por comparación y taxonomía, 5- el análisis estadístico de las regularidades de las poblaciones y el cálculo de probabilidades, y 6- la derivación histórica del desarrollo genético.

Hacking resume su posición, que define como “anarco-racionalista” y como no subjetivista:

1. Hay diferentes estilos de razonamiento, con sus propias emergencias y trayectorias históricas; algunos han perecido, otros perduran y en adelante podrían emerger nuevos;
2. Las proposiciones adquieren “positividad” –son verdaderas o falsas– sólo dentro de su estilo de razonamiento;
3. Las categorías de lo que puede ser verdadero o falso son históricas, con relación al desarrollo contingente de ciertos estilos de razonamiento;
4. En nuestra tradición han emergido diversas categorías de lo que puede ser verdadero o falso;
5. No podemos juzgar si un sistema de razonamiento alternativo es mejor o peor que el nuestro, porque las proposiciones con las cuales razonamos no existen con independencia del estilo de razonamiento correspondiente.

Más tarde, en Hacking (1992), para diferenciarse de otras conceptualizaciones de “estilo” como la de Gayon (1996), Hacking recurre a Fleck (edición en español: Fleck, 1987); quien, más allá de divergencias, le sirve para resaltar el carácter impersonal y colectivo del estilo de razonamiento, que marca los límites de lo que se puede decir. En este mismo sentido, menciona “las epistemes y las formaciones discursivas de Michel Foucault”¹⁰. También en este texto, Hacking subraya que elige el término “razonamiento” en vez del crombiano “pensamiento” para enfatizar su aspecto no sólo mental y privado, sino sobre todo público; por otra parte, reconoce que “razonamiento” no alcanza para referirse a los aspectos manuales y artesanales de la actividad científica. Pero fundamentalmente elige este término por sus resonancias kantianas, y en tal sentido expresa: “Mi estudio es una continuación del proyecto de Kant de explicar cómo es posible la objetividad” (Hacking, 1992, p. 4)¹¹. Continuación que, sin embargo, procura pensar lo que Kant no pensó: que la razón científica es un producto histórico y colectivo, y que lo que llamamos “objetividad” en ciencias ha sido establecido por estilos de razonamiento científico; los cuales, habiéndose constituido “a través de interacciones y negociaciones microsociales”, se han tornado autónomos y han llegado a ser los estándares de objetividad en una materia determinada.

A diferencia de Crombie –historiador con intereses en una antropología histórica comparada– *Hacking filósofo usa a Crombie para una “historia del presente” à la Foucault*; historia que nos permita ver los seis estilos hoy, tal y como perviven y definen lo que es

¹⁰ Al respecto, resulta valioso el trabajo realizado por Vagelli (2012a), quien señala los puntos de contacto entre las “formaciones discursivas” de Foucault y los “estilos de razonamiento científico” de Hacking.

¹¹ “My study is a continuation of Kant’s project of explaining how objectivity is possible” (Hacking, 1992, p. 4).

importante para nosotros aquí y ahora. Asimismo como filósofo, Hacking procura determinar criterios o condiciones necesarias para considerar a algo como un estilo, y en ese sentido propone dos:

1. Cada estilo debe introducir novedades (algunas o todas de éstas): nuevos tipos de objetos, nuevos tipos de debates ontológicos, nuevas condiciones de enunciación de lo verdadero, nuevas clases de sentencias que adquieren positividad, nuevas leyes, nuevas modalidades de medida, nuevos tipos de clasificación y de explicación;
2. Cada estilo es auto-autenticante y persiste en el tiempo gracias a sus propias técnicas de auto-estabilización (siendo algunas técnicas estabilizantes más eficaces que otras).

En Hacking (2002a) se detiene especialmente en el estilo taxonómico, en las clasificaciones y en el “efecto bucle” de las clases humanas y concluye preguntándose por la idea de naturaleza presupuesta en las clasificaciones de las ciencias naturales. En este contexto, refiriéndose a Nietzsche y a Foucault, subraya la expresión “ontología histórica”, que define su perspectiva filosófica.

En la Introducción de Hacking (2002b), éste se autodenomina un “nominalista dinámico”/ un “realista dialéctico” y agrupa sus diversas indagaciones e intereses bajo la rúbrica “ontología histórica”. Afirma que ésta es “una generalización” de la que propusiera Foucault, pues su intento apunta a dar cuenta del “llegar a ser”, no sólo de nosotros mismos, sino también del devenir de toda clase de objetos –podríamos decir, de toda clase de objetos vinculados con los estilos de razonamiento científico. A la vez, reconoce que su ontología histórica es “más limitada” que la de Foucault, pues a pesar de tratar de “preservar sus tres ejes” (saber, poder y ética), se vincula más con sus intentos arqueológicos que con aquéllos genealógicos. En su versión más audaz, la ontología histórica apunta a entender y a resolver problemas presentes; en su versión más modesta –en la que Hacking ubicaría su propio intento– es análisis conceptual, un análisis de los conceptos como “palabras en sus sitios” (las sentencias en las cuales la palabra es realmente usada, quiénes las pronuncian, con qué autoridad, en qué instituciones, para influenciar a quién, con qué consecuencias...). Análisis conceptual que “hace usos” de la historia, pero que no es historia sino “un modo de hacer filosofía”.

Acerca de algunas “mezclas” de historia y filosofía, Hacking se refiere especialmente a los trabajos de Daston y de otros investigadores del MPIWG, así como al de Davidson -de la Universidad de Chicago-, todos ellos etiquetados como “HE”. Como lo expresara anteriormente en Hacking (1999), considera más adecuado denominar a aquellos trabajos como “meta-epistemología histórica”, pues estudian las trayectorias históricas de conceptos muy generales como los de objetividad, racionalidad, hecho, etc., que organizan el campo o el “espacio de posibilidades” de lo que conocemos y creemos y de lo que pensamos acerca de ello. Así entendida, esta meta-epistemología histórica sería una parte de la ontología histórica.

A partir de Hacking (2010), abandona la denominación “estilos de razonamiento científico” y elige otra para nombrar su proyecto estilos: “estilos de pensamiento & acción

científicos” –“*styles of scientific thinking & doing*”¹². En este texto, distingue este concepto de los conceptos interrelacionados “estilo de pensamiento” y “colectivo de pensamiento” de Fleck; mientras que el suyo es una herramienta filosófica, los conceptos fleckianos son herramientas sociológicas. Sin embargo, pueden plantearse relaciones entre ellos: un “estilo de pensamiento” característico de un “colectivo de pensamiento” puede comenzar en un tiempo y en un lugar y crecer al interior de un estilo de pensamiento & acción científicos, estabilizarse y continuar hasta el presente; de este modo, algo inicialmente local puede llegar a ser patrimonio común de la humanidad.

Hacking (2012) constituye la última una puesta al día de su proyecto estilos, después de treinta años de su primera formulación y tras las críticas que le efectuara Kusch (2010)¹³. En esta revisión, Hacking vuelve a denominar su proyecto como “estilos de pensamiento & y acción científicos”. Reconoce que, cuando escuchó exponer a Crombie acerca de los estilos en 1978, estaba bajo el influjo de “las tempranas arqueologías” de Foucault e inspirado por el texto *Contra el Método* de Feyerabend. Desde entonces, ha venido usando la ya clásica lista de Crombie, pero convirtiéndola en una herramienta para su proyecto *filosófico*. El mismo consiste en “usar el pasado para entender el presente”, en mostrar las diferentes trayectorias históricas de cada uno de estos métodos de investigación para entender cómo estos devenires (concebidos como ciclos de “contingencias sin fin” de emergencias, cristalizaciones, sedimentaciones y estabilizaciones) han ido moldeando la razón y los criterios y cánones de racionalidad, objetividad y verdad que perduran hasta el presente. Aclara también que su proyecto estilos no tiene ya compromisos con la palabra “estilo”, usada de varias maneras particularmente en historia del arte. Constata que su denominación “estilos de razonamiento científico” fue usada desde sus inicios en muy diferentes contextos, sin relación alguna con su propuesta. Nos resulta interesante la primer nota al pie de página de este texto; allí Hacking menciona el subtítulo del libro de Davidson (2001), infiere que este subtítulo pudo haber promovido la asociación de su proyecto estilos con la HE y se preocupa en aclarar que él nunca ha pensado su proyecto dentro de la HE¹⁴. Aquí como en otros sitios, Hacking manifiesta sus diferencias y distancias respecto de la HE.

Asimismo en este texto, Hacking renuncia ya a la pretensión filosófica de determinar “condiciones necesarias y suficientes” para definir a algo como estilo. Las razones son ecológicas más que lógicas o pragmáticas: los seis estilos de Crombie son los que han persistido y han llegado a ser nuestros estándares de racionalidad, “son los que usamos”. La metáfora elegida por Hacking para dar cuenta de la emergencia de un nuevo estilo es la de “cristalización”, que le permite identificar rupturas y emergencias de lo nuevo en un conjunto de prácticas históricamente situadas. Desde esta perspectiva, Hacking aprovecha aportes historiográficos contemporáneos para localizar cristalizaciones –como los de Netz referidos al descubrimiento de la prueba demostrativa

¹² Transcribimos esta expresión en español con el símbolo “&” (et o ampersand), porque precisamente en Hacking (2010) manifiesta que intenta conservar dicho símbolo como su propia marca registrada, que lo diferencia de otros usos que se han hecho su anterior expresión “estilos de razonamiento científico” (cf. Hacking, 2010, p. 230, nota 5).

¹³ Cabe notar que Kusch (2010) considera que los desarrollos teóricos de Hacking merecen el título de “Historical Epistemology”. Asimismo, en este artículo, Kusch propone tres tesis críticas: 1) Hacking adopta “una posición historiográfica anticuada” debido a su relación con Crombie, 2) Hacking no logra distanciar la HE del relativismo epistémico y 3) Hacking no ha brindado aún un “criterio convincente” para individualizar los estilos de razonamiento.

¹⁴ “For the record: I cannot recall ever thinking of the styles project as filed under ‘historical epistemology’” (Hacking, 2012, p. 601, nota 1).

en matemática y los de Shapin y Schaffer referidos a la exploración experimental– y rectifica su idea de que el “estilo laboratorio” era un nuevo estilo –entendiendo ahora que se trata más bien de una cristalización dentro del segundo estilo. La idea de “cristalización” encaja bien con una concepción de la verdad y la veracidad que sostiene que han habido desplazamientos en el decir la verdad acerca de algo (X) en un dominio de investigación, en una época determinada (Y), en conexión con una figura emblemática (Z). El proyecto estilos queda abierto; entre otros tópicos de interés, Hacking invita a seguir pensando la “personalidad” de cada estilo: cada uno tiene una diferente relación con la verdad –particularmente, los tres segundos estilos respecto de los tres primeros– y cada uno tiene diversos modos de introducir nuevas clases de objetos en el discurso científico.

Podemos apreciar que el concepto hackiniano de estilos es clave en su propuesta filosófica. Rescatamos dos valoraciones de este concepto, que nos invitar a pensar en su lugar, función y sentido al interior del proyecto de Hacking, así como en su utilidad como herramienta teórica. En primer lugar, referimos a Braunstein (2008), quien considera que el concepto de estilo de pensamiento científico, elaborado entre otros por Hacking, permite superar ciertas controversias clásicas en historia de las ciencias, tales como la oposición entre historia filosófica de las ciencias y estudios sociológicos locales de la ciencia, la querrela entre continuismo y discontinuismo, así como los debates tradicionales sobre el realismo y la ontología. En segundo lugar, Vagelli (2012a) subraya que el concepto de estilos de razonamiento científico de Hacking es “un concepto de transición” entre la fase originaria y la fase contemporánea de la EH/HE, “una piedra de toque” en relación con otras nociones (“paradigma” de Kuhn, “esquema conceptual” de Quine y “formaciones discursivas” de Foucault) y en definitiva, “el epicentro de su posición epistemológica” y por ende, “un punto de apoyo privilegiado para extraer la originalidad de su versión de la EH”.

4. La cuestión de la Epistemología Histórica como estilo epistemológico

Abordamos aquí la pregunta planteada al inicio para intentar responder si la EH/HE puede ser entendida como un estilo epistemológico. Como Hacking (2002) lo señala, hay muchos modos de hacer filosofía¹⁵. Ahora bien, los diversos modos de hacer filosofía, particularmente epistemología, ¿pueden ser pensados como *estilos*, especialmente en el caso de la EH/HE?

En el intento por responder esta compleja cuestión, nos iluminan los análisis de Vagelli (2012b), quien se pregunta si el concepto de estilo, tal como ha sido rectificado por Hacking en términos de “estilos de razonamiento científico”, puede ser usado como categoría historiográfica para una historia de la filosofía y particularmente, en un movimiento recurrente, para una historia de la EH/HE. Su examen concluye mostrando que, dentro de los límites del concepto elaborado por Hacking, “se torna complicada” la extensión del mismo “a la historia de la filosofía o de la epistemología”, siendo sólo un concepto “de segundo orden” para la historia de las ciencias. Tampoco dentro de los límites de este concepto hackiniano podría decirse que la epistemología, y aún la EH/HE, introducen “novedades” del tipo de las que describe Hacking, sino más bien un conjunto de categorías historiográficas, entre las que encontramos la propia noción de estilo. A la luz de sus elaboraciones posteriores (Vagelli, comunicación personal, 8 de diciembre de 2015), podríamos afirmar que Vagelli realiza en un primer momento un análisis de los

¹⁵ “There are many ways to do philosophy” (Hacking, 2002, p. 25).

desarrollos y de los usos del concepto de estilo en Hacking como un preliminar ineludible para luego intentar superar los límites de este concepto y extender su uso a la historia de la filosofía, de la epistemología y de la EH.

Estos análisis nos invitan a volver sobre los textos de Hacking para interrogarnos primeramente por qué *no* podría usarse la noción hackiniana de estilos en filosofía y epistemología y en sus historias –particularmente, en la historia de la EH/HE-, para luego preguntarnos por qué *sí* sería posible y provechoso hacerlo.

En ese sentido, ante todo hay que tener en cuenta que Hacking afirma que su doctrina de los estilos de razonamiento auto-autenticantes es una estrategia que busca explicar “lo que es peculiar de la ciencia”, distinguiéndola de otras indagaciones (como las humanísticas y éticas) (cf. Hacking, 2002b, p. 196). Podría decirse entonces que, a diferencia de las ciencias, las indagaciones humanísticas, éticas en particular y filosóficas en general, no constituyen estilos de razonamientos auto-autenticantes, razón por la cual los estilos hackinianos no servirían como estrategia para dar cuenta de lo peculiar de la filosofía. A partir de estas afirmaciones, se sigue que *no* podría usarse la noción hackiniana de estilos para dar cuenta, ni de lo peculiar de la filosofía/epistemología, ni de su historia, particularmente la historia de la EH/HE. Sin embargo, a pesar de esta restricción, podemos pensar en una de las ventajas de sostener este *no*: al distinguir lo peculiar de los estilos de pensamiento & acción científicos de lo peculiar de la filosofía/epistemología, así como al reservar la categoría hackiniana de estilos sólo para una historia de las ciencias, se preservaría a la epistemología como una actividad de análisis, crítica y reflexión acerca de los estilos de pensamiento & acción científicos y acerca de los usos de esta noción en diferentes contextos. Esta preservación es de importancia decisiva, particularmente ante diversas tendencias y proyectos actuales de “naturalización de la epistemología” –algunos de los cuales podrían tratar de encontrar respaldo en algunas indicaciones recientes del propio Hacking, cuando se refiere a los desarrollos de la “historia cognitiva” y de la “historia ecológica –(cf. Hacking, 2012; Vagelli, 2014b).

Por otra parte, veamos por qué *sí* podría usarse la noción de estilos en filosofía y en epistemología, y con qué provecho. Como hemos referido, Braunstein extiende el uso del concepto de estilos de razonamiento científico para dar cuenta de los rasgos comunes de la epistemología y de la historia de las ciencias francesas, sin perder de vista las diferencias entre sus principales exponentes. Acerca de este uso del concepto en Braunstein, Vagelli (2012b) expresa que esta “torsión” del concepto de estilo abre perspectivas interesantes y asimismo, valora que Braunstein propuso una aplicación filosófica, particularmente epistemológica, del estilo como concepto analítico. Habría pues, un uso posible de la noción de estilo en epistemología.

Resulta aquí oportuno subrayar que Hacking señala que la palabra “estilos” puede ser usada para describir otro tipo de prácticas además de las seis propuestas por Crombie: “Estilos, géneros, indagaciones y modos de descubrir: todas esas palabras pueden ser usadas para describir muchas más clases de prácticas que las seis de Crombie (Hacking, 2012, p. 601)¹⁶. Nos preguntamos entonces si esta afirmación de Hacking abre la posibilidad de describir como estilos no sólo a prácticas de nuevos estilos de razonamiento & acción científicos, sino también a prácticas de filosofía/epistemología. Estimamos que sí.

¹⁶ “Styles, genres, inquiry and ways of finding out: all these words can be used to describe a great many more kinds of practices than Crombie’s six” (Hacking, 2012, p 601).

Recordemos además que en Hacking (2012) renuncia a la pretensión filosófica de determinar “condiciones necesarias y suficientes” para definir a algo como estilo de razonamiento & acción científicos. Esto abriría la posibilidad de otros usos de la palabra estilo, aun cuando no se cumplimentaran todas y cada una de las condiciones antes mencionadas.

Asimismo, al referirse a la distinción propuesta por Williams (2002) entre “verdad” y “veracidad” –“*truth*” y “*truthfulness*”–, Hacking (2012) afirma que los filósofos “hacen” distinciones para determinados propósitos. Puede entonces pensarse que lo propio de la filosofía es hacer distinciones. Si esto es así, entonces la filosofía también crea, trae al ser novedades, puede derrocar lo establecido al crear un nuevo nombrar y unas nuevas distinciones. Aquí cabe recordar aquella afirmación en Hacking (2002a), acerca de la ontología de Foucault: “Michel Foucault hizo mucho más que ontología histórica. Ayudó a traer cosas al ser. La suya fue tanto una ontología creativa como una histórica” (Hacking, 2002a, p. 8)¹⁷. Si algunos filósofos, como ejemplarmente lo hizo Foucault, pueden ayudar a traer cosas al ser, si existe la posibilidad de que se genere una ontología creativa, entonces dicha ontología cumple al menos con una de las condiciones de lo que es un estilo.

Afirmar que la filosofía / epistemología sí es un estilo sería provechoso para reconocer lúcida y críticamente la pertenencia a un determinado modo de hacer filosofía/ epistemología en el cual hemos sido formados, para promover el respeto por otros modos de hacer filosofía/ epistemología y para plantear una comunicación posible entre diversos modos. Por ejemplo, podemos haber sido formados en un modo de hacer filosofía a partir de problemas. Ahora bien, el reconocimiento lúcido y crítico de la pertenencia a este modo de filosofar exige un trabajo de análisis. Al respecto, citamos a Hacking: “‘el problema’ como un modo definitivo de filosofar fue canonizado en inglés alrededor de 1910 con títulos de G.E. Moore, Williams James y Bertrand Russell” (Hacking, 2002b, p. 12)¹⁸. Los modos de filosofar tienen su historia, llegan a ser modos canónicos en una lengua, un momento y un lugar, asociado a nombres emblemáticos. Esta cita nos hace pensar en la posibilidad de un análisis conceptual, necesariamente histórico, de la emergencia, establecimiento y “canonización” de determinados modos hacer filosofía/epistemología, como exigencia ineludible para la revisión del propio modo, el respeto de otros y la posibilidad de comunicación y entendimiento entre diversos, aprendiendo “a razonar de un nuevo modo” (cf. Hacking, 2002b, p. 171).

En síntesis, podemos entonces afirmar que sí cabe hablar de estilos filosóficos (y particularmente epistemológicos), pero no en el sentido de los estilos de pensamiento & acción científicos propuestos por Hacking. Asimismo, pensamos que resulta fecundo plantear relaciones entre ambos estilos. Particularmente, nos interesan las relaciones entre lo que denominaríamos el “estilo” de la epistemología histórica en su fase originaria y los estilos de pensamiento & acción científicos propuestos por Hacking.

A partir de los desarrollos de Hacking, podemos entonces esbozar las siguientes notas respecto de las relaciones entre estilos epistemológicos y estilos de pensamiento & acción científicos:

¹⁷ “Michel Foucault did far more than historical ontology. He helped bring things into being. His was a creative ontology as well as a historical one” (Hacking, 2002a, p. 8).

¹⁸ “‘the problem’ as definitive of a mode of philosophizing was canonized in English around 1910 with titles by G.E. Moore, William James, and Bertrand Russell” (Hacking, 2002b, p. 12).

- Los estilos de pensamiento & acción científicos introducen nuevos debates ontológicos. Hay, pues, un filosofar en el espacio abierto por estos estilos. Cabe recuperar aquí aquel rasgo distintivo del estilo francés en epistemología que Braunstein presenta como “una filosofía contemporánea de las ciencias” (cf. Braunstein, 2002, p. 6). Una epistemología de este estilo, que presta interés a la historia y a la actualidad de las ciencias, procurará situar los debates ontológicos en relación a los nuevos tipos de objetos que un estilo de pensamiento & acción científico ha introducido. Ello exigirá una epistemología que, en términos de Bachelard, puede caracterizarse como una epistemología pluralista, coherente, “del detalle”, atenta a la actividad de una región científica contemporánea y a las novedades que en ella emergen.
- Cuando se localizan y se caracterizan los eventos que marcaron radicales innovaciones en la historia de las ciencias y cuando se asocian a ellos los héroes o emblemas de cada estilo de pensamiento & acción científicos, se pone en juego un estilo epistemológico que determina la interpretación de tales eventos y de tales héroes o emblemas. Aquí podemos proponer como ejemplo lo señalado por Hacking (2012): Galileo ha sido emblema de dos estilos, el experimental y el de modelización hipotética, y ha sido presentado de diversas maneras: ya como el “Galileo experimentador” de Kant y de Drake, ya como el Galileo de la matematización de la naturaleza de Koyré y de Husserl. Será parte del ejercicio epistemológico indagar y exponer en detalle cómo estas interpretaciones diversas han estado determinadas por diferentes estilos epistemológicos.
- También se ponen en juego estilos epistemológicos cuando se ubican y se interpretan continuidades y/o rupturas en la historia de las ciencias. Por ejemplo, en el caso del propio Hacking, podría afirmarse que desde un estilo epistemológico sensible a rupturas y a continuidades, pero más profundamente interesado en los aspectos “braudelianos” del conocimiento científico (Hacking, 1999b), Hacking (2012) elige la metáfora de las cristalizaciones para dar cuenta de la emergencia de los estilos y apela a técnicas auto-estabilizantes y a formas de auto-justificación y de auto-autenticación de los estilos de razonamiento (Hacking, 1992 y 1999b) para dar cuenta de la permanencia y de la larga duración de los mismos.
- Las relaciones entre modos de hacer filosofía/ epistemología y los estilos de pensamiento & acción científicos remiten, en definitiva, a las relaciones entre filosofía/epistemología y ciencias. Proponemos volver a pensar estas relaciones a partir de una recuperación de algunas tesis fundamentales de representantes de la etapa originaria de la EH, para llegar con solidez hasta Hacking. Nos ilumina aquí particularmente Canguilhem, para quien “La filosofía es una reflexión para la cual toda materia extranjera es buena, y estaríamos dispuestos a decir, para la cual toda buena materia tiene que ser extranjera” (Canguilhem, 1971, p. 11). Las materias extranjeras que promovieron la reflexión filosófica de Canguilhem fueron la medicina como arte o técnica clínica –a la que acudió para acercarse a “problemas humanos concretos”–, y de manera más amplia, las ciencias de la vida, interesándose en el nacimiento, formación, importaciones y exportaciones de determinados conceptos de funciones biológicas, al interior del conjunto de prácticas técnicas, políticas e ideológicas que los hicieron posibles, en relación con intenciones, decisiones y

proyectos normativos. Ahora bien, para Canguilhem (1993), sólo puede hablarse de verdad y de objetividad en el dominio de la ciencia. Pero ello no significa que la filosofía no tenga objeto (cf. Le Blanc, 2004, p. 17). Precisamente, la filosofía es una reflexión que vuelve a abrir problemas que para la ciencia estaban resueltos (Canguilhem, 1971, p. 13), es evaluación de valores –particularmente de los valores de verdad y de no-verdad en los discursos científicos–, es interrogación acerca de la finalidad de la verdad científica, aspira a una “buena” totalización de la experiencia de una época, a difundirse en distintos campos de la cultura y a tener impactos en nuestra existencia cotidiana (cf. Canguilhem, 1993). En relación con nuestro tema, podría decirse que este “estilo de pensamiento” –así lo ha llamado Macherey (1996)– promueve una reflexión epistemológica e histórica de las ciencias de la vida y un compromiso con nuestra actualidad, poniendo el acento en la filosofía como el lugar donde la verdad de la ciencia se confronta con otros valores, como los estéticos, éticos (Canguilhem, 1993) y/o políticos, atenta a todo aquello que hoy afecta a la vida humana en su entorno y a lo que se puede crear, a partir de la normatividad vital y social propia de la vida humana.

5. Consideraciones finales

Nos parece que lo que está en juego al plantear la cuestión del estilo filosófico y su relación con los estilos de razonamiento & acción científicos propuestos por Hacking es, en definitiva, el sentido mismo de la filosofía y de la epistemología.

En su reciente disertación de doctorado, Vagelli (comunicación personal, 4 de diciembre de 2015), entre otras cuestiones, plantea relaciones entre estilos de pensar y estilos de existencia en Foucault, Hacking y Davidson y, superando los límites del concepto hackiniano de estilo, afirma la posibilidad de aplicar este concepto a la historia de la filosofía y de la epistemología. A la espera de la lectura de esta disertación y de su difusión en nuestro medio, nos preguntamos si pueden recuperarse algunas cuestiones planteadas por representantes de la fase inicial u originaria de la EH, que nos pueden aproximar a la idea de la epistemología como *estilo de pensar* y como *trabajo sobre sí*. En tal sentido,

- De Bachelard, recuperamos la propuesta de que la epistemología es no sólo una actividad atenta a la historia de las ciencias y a su actualidad, sino fundamentalmente un trabajo sobre sí que modifica al sujeto de conocimiento científico. En tal sentido, la epistemología tiene que ver con una “ontología discursiva” (Bachelard, 2004) en la que se van dando conjuntamente procesos de construcción y rectificación objetiva y procesos de formación e incesante reforma subjetiva, en las dialécticas entre obstáculos y actos epistemológicos, a través de ejercicios de psicoanálisis del conocimiento objetivo (Bachelard, 1991) y de vigilancia epistemológica (Bachelard, 1978). Este estilo implica la constitución de una subjetividad no sólo en el aspecto epistemológico, sino también ontológico, ético y estético (Becerra Batán, 2010).
- De Foucault, recuperamos su concepción de la filosofía como práctica de cuidado de sí, con sus formas de reflexividad y prácticas de subjetividad (Foucault, 2002), la filosofía como trabajo crítico del pensamiento sobre sí mismo, como ensayo, ascesis o prueba modificadora de sí mismo en el juego

de la verdad (Foucault, 1984), para pensar y hacer de otro modo. Coincidimos aquí con Paltrinieri (2012): la obra de Foucault nos obliga a ir hacia el exterior de la filosofía, “a hacer de la filosofía no una ‘arquitectura de sistemas’, sino más bien una práctica, un estilo de reflexión y tal vez de existencia” (Paltrinieri, 2012, p. 277)¹⁹.

Hacking se vincula con estos referentes de la EH, especialmente con Foucault. Podría decirse que *su peculiar modo de hacer filosofía supone un modo de hacer epistemología que es una epistemología histórica* (aun cuando tenga reservas con esta denominación). Y que *este modo de hacer filosofía/ epistemología implica un modo de hacerse a sí mismo*. En tal sentido, Hacking reconoce el impacto de Foucault en su propia formación y lo que ha llegado a ser: un filósofo de tradición analítica “con sensibilidad histórica”, cuyo modo de hacer filosofía consiste en “echar un vistazo” (“*take a look*”) y en “deshacer problemas” (“*undoing problems*”) –cf. Ribeiro Mendes, 2012. Subrayamos que este “echar un vistazo” exige un trabajo sobre sí para mirar con nuevos ojos la rica complejidad del mundo donde nos hallamos pensando. Asimismo, este trabajo sobre sí permite “deshacer problemas” –que es a la vez un deshacerse de aquel modo canónico de hacer filosofía centrado en problemas– mediante un análisis conceptual que ya no puede dejar de hacer usos de la historia; de este modo, se trata ahora de un análisis de los conceptos como palabras en sus sitios *históricos*, para entender nuestra actualidad e intervenir en ella.

La epistemología histórica puede ser entonces concebida y practicada como un estilo, un ejercicio de análisis, crítica y reflexión de lo que han implicado e implican hoy, para la vida humana en el mundo, los estilos de pensamiento & acción científicos; estilo orientado hacia la transformación de las prácticas científicas y de nosotros mismos, desde nuestra actualidad.

6. Referencias

- AAVV. (2012). *Conference Epistemology and History. From Bachelard and Canguilhem to Today's History of Science*. Preprint 434, Berlin, Alemania, Max Planck Institute für Wissenschaftsgeschichte.
- Almeida, T. S. & Camolezi, M. (2016). Entretien avec Jean-François Braunstein. *Intelligere, Revista de História Intelectual*, São Paulo, Brasil, v. 2, n. 1 [2], 2016, p. 171-186.
- Bachelard, G. (2004). Ontología discursiva. En G. Bachelard, *Estudios*, Buenos Aires, Argentina: Amorrortu, p. 99-110.
- Bachelard, G. (1991). *La formación del espíritu científico*, D.F., México: Siglo XXI Editores.
- Bachelard, G. (1978). *El racionalismo aplicado*, Buenos Aires, Argentina: Paidós.
- Becerra Batán, M. (2010). *Epistemología y subjetividad en Gaston Bachelard*. Tesis de Doctorado en Filosofía, FFyH, UNC, Córdoba, Argentina, ejemplar de Biblioteca FFyH.
- Braunstein, J-F. (2002). Bachelard, Canguilhem, Foucault. Le ‘style français’ en épistémologie. En P. Wagner (dir.) (2002) *Les philosophes et la science*, Paris, Francia: Gallimard.
- Braunstein, J-F. (2006). Fleck, Canguilhem, Foucault. Ludwig Fleck et le ‘style français’ en philosophie des sciences. En J. Fer.; N. Jas & I. Löwy, (Eds), *Penser avec Fleck* –

¹⁹ “À faire de la philosophie non point une ‘architecture des systèmes’, mais plutôt une pratique, une style de réflexion et peut-être d’existence” (Paltrinieri, 2012, p. 277)

- Investigating a Life Studying Life Sciences*, Zürich, Suiza: Collegium Helveticum Heft 7, p. 91-100.
- Braunstein, J-F. (2008). *L'histoire des sciences. Méthodes, styles et controverses*, Paris, Francia: Vrin.
- Braunstein, J-F. (2012). Historical Epistemology. Old and New. En *AAVV Conference Epistemology and History. From Bachelard and Canguilhem to Today's History of Science*. Preprint 434, Berlin, Alemania, Max Planck Institute für Wissenschaftsgeschichte, p. 33-40.
- Canguilhem, G. (1971). *Lo normal y lo patológico*. México: Siglo XXI Editores.
- Canguilhem, G. (1968). L'histoire des sciences dans l'oeuvre épistémologique de Gaston Bachelard. En G. Canguilhem, *Études d'histoire et de philosophie des sciences*. Paris: Vrin, p. 173-186.
- Canguilhem, G. et al. (1993). *Philosophie et vérité*. Georges Canguilhem, Michel Foucault, Jean Hypolite, Paul Ricoeur, Alain Badiou y Dina Dreyfuss. Émission produite par la Radio-Télévision scolaire y CNDP, 1965. Paris: CNDP y Ed. Nathan.
- Daston, L. (2008). Une histoire de l'objectivité scientifique. En J-F. Braunstein (dir.) *L'histoire des sciences. Méthodes, styles et controverses*, Paris, Francia: Vrin, p. 363-375.
- Davidson, A. (2001). *The emergence of sexuality: Historical Epistemology and the formation of concepts*, Cambridge, Mass, EEUU: Harvard University Press.
- Feest, U. & Sturm, T. (2011). What (good) is Historical Epistemology? Editor's Introduction. *Erkenntnis*, 75/3 (2011) Special Issue, p. 285-302.
- Fleck, L. (1987). La génesis y el desarrollo de un hecho científico. Introducción a la teoría del estilo de pensamiento y del colectivo de pensamiento, Madrid, España: Alianza Editorial.
- Foucault, M. (1969). *Archéologie du savoir*, Paris, France: Gallimard.
- Foucault, M. (1984). *Histoire de la sexualité. 2- L'usage des plaisirs*, Paris, France: Gallimard.
- Foucault, M. (2002). *La hermenéutica del sujeto. Curso en el Collège de France (1981-1982)*, Buenos Aires, Argentina: Fondo de Cultura Económica de Argentina.
- Gayon, J. (2006). Bachelard y la historia de las ciencias. En J-J. Wunenburger, *Bachelard y la epistemología francesa*, Buenos Aires, Argentina: Nueva Visión, p. 39-86.
- Gayon, J. (1996). De la catégorie de style en histoire des sciences. *Alliage*, numéro 26, printemps 1996.
- Gingras, I. (2010). Naming without necessity: on the genealogy and uses of the label "historical epistemology", *Note de recherche 2010-01*, Montreal, Canada: CIRST.
- Hacking, I. (1982). Language, Truth and Reason. En Hollis & Luke. *Rationality and relativism*, MIT Press, p. 48-66.
- Hacking, I. (1990). *The taming of chance*, Cambridge, EEUU: Cambridge University Press.
- Hacking, I. (1992). Style for historians and philosophers. *Studies in History and Philosophy of Science*. Vol 23 (1), p. 1-20, 1992.
- Hacking, I. (1999). Historical meta-epistemology. En W. Carl & L. Daston (eds.), *Wahrheit und Geschichte. Abhandlungen der Akademie der Wissenschaften in Göttingen*, Göttingen: Vanderhoed & Ruprecht, p. 53-77.
- Hacking, I. (1999b). La autojustificación de las ciencias de laboratorio. En Ambroggi, A. (Editor) *Filosofía de la ciencia. El giro naturalista*, Palma, España: Universitat de les Illes Balears, p. 213-250.

- Hacking, I. (2002a). Inaugural Lecture: Chair of Philosophy and History of Scientific Concepts at the Collège de France, 16/2/2001. *Economy and Society*, Vol. 31, N° 1, p. 1-14.
- Hacking, I. (2002b). *Historical Ontology*, Cambridge, EEUU: Harvard University Press.
- Hacking, I. (2010). Lloyd, Daston, Nurture and 'Style'. *Interdisciplinary Science Reviews*, Vol 23, N° 3-4 (2010), Institute of Materials, Minerals and Mining, Maney Publishing, p. 231-240.
- Hacking, I. (2012). 'Language, Truth and Reason' 30 years later. *Studies in History and Philosophy of Science*. Vol 43 (2012), p. 599-609.
- Kusch, M. (2010). Hacking's Historical Epistemology. A critique of styles of reasoning. *Studies in History and Philosophy of Science*. Vol 41 (2010), p. 158-173.
- Le Blanc, G (2004). *Canguilhem y las normas*, Buenos Aires, Argentina: Nueva Visión
- Lecourt, D. (1969). *L'épistémologie historique de Gaston Bachelard*. Paris, France: Vrin.
- Lecourt, D. (2007). Georges Canguilhem, le philosophe. En J-F. Braunstein (ed). *Canguilhem. Histoire des sciences et politiques du vivant*, Paris, France: PUF.
- Macherey, P. (1996). Georges Canguilhem : un style de pensée. *Cahiers philosophiques*, N° 69, décembre 1996, Paris, France: CNDP, p. 47-56.
- Méthot, P. O. (2012). On the genealogy of concepts and experimental practices. Rethinking Georges Canguilhem's Historical Epistemology. En *Conference Epistemology and History. From Bachelard and Canguilhem to Today's History of Science*. Preprint 434, Berlin, Alemania, Max Planck Institute für Wissenschaftsgeschichte, p. 117-144.
- Paltrinieri, L. (2012). *L'expérience du concept. Michel Foucault entre épistémologie et histoire*, Paris, France: Publications de la Sorbonne. La Philosophie à l'œuvre.
- Rheinberger, H-J. (2010). *On historicizing Epistemology, An Essay*, Stanford CA, EEUU: Stanford University Press.
- Ribeiro Mendes, J. (2012). Um modo de fazer filosofia: a ontologia histórica de Ian Hacking. *Diacrítica. Serie Filosofia e Cultura*. N° 26/2, 2012. Universidade do Minho, Braga, Portugal, p. 105-120.
- Vagelli, M. (2012a). Discours et styles de raisonnement: Hacking face à Foucault. Documento inédito. Ateliers Doctoreaux Philosophies Contemporaines, Paris, France, 9/3/13.
- Vagelli, M. (2012b). La notion de style en histoire des sciences, est-elle applicable à l'histoire de l'épistémologie? Documento inédito. Séminaire « Archives ». La Philosophie comme objet. Sociologie et Histoire de la Philosophie, CIEPFC Paris, France, 14/12/12.
- Vagelli, M. (2014a). *Journées d'études: Épistémologie historique: commencements et enjeux actuels*. Université Paris 1, Panthéon Sorbonne. Appel à contributions/Call for papers, <https://episthist.hypotheses.org/>
- Vagelli, M. (2014b). Ian Hacking. The Philosopher of the Present. An interview by Matteo Vagelli. *Iride* a. XXVII, n 72, maggio-agosto 2014 / *Iride*, v. 27, issue 72, May-August 2014, p. 239-269.
- Vagelli, M. (2015). *Deuxième Journées d'études: Épistémologie historique*. Université Paris 1, Panthéon Sorbonne. Appel à contributions / Call for papers, <https://episthist.hypotheses.org/>
- Williams, B. (2002). *Truth and truthfulness: an essay in genealogy*, Princeton, EEUU: Princeton University Press.

Actuar intencionalmente: esbozo de una teoría de la intención

Diego Lawler¹

Recibido: 28 de julio de 2016.

Aceptado: 2 de septiembre de 2016.

Resumen. En este trabajo ofrezco un esbozo de una teoría de la intención. Propondré que actuar intencionalmente es, en términos generales, actuar según una intención. Sin embargo, el análisis completo de esta propuesta traerá consigo la discusión sobre qué se entiende por la noción de intención. Esto último depende, en cierto sentido, de cómo se explique el vínculo que existe entre intentar hacer algo y hacer ese algo intencionalmente o con cierta intención. En los dos apartados siguientes discutiré la teoría de la intención. Luego consideraré brevemente dos problemas relacionados con la idea de intención elucidada desde el fenómeno de los planes y la planificación. A continuación analizaré las intenciones sobre el futuro en su condición de insumos (*inputs*) y resultados (*outputs*) de razonamientos prácticos. Después bosquejaré la estructura básica de la acción intencional haciendo hincapié en la idea de plan de acción. Finalmente, consideraré un problema filosófico ligado a la idea de que la intención sobre el futuro constituye el caso central de intención: la intención en la acción o el caso de la acción intencional espontánea.

Palabras clave: acción intencional – intención – planes – razonamiento práctico.

Título: Acting intentionally: a sketch of a theory of intention.

Abstract. In this paper I offer a sketch of a theory of intention. Generally speaking, I propose that acting intentionally is acting with an intention. However, a complete analysis of this proposal will involve the discussion of what is meant by the notion of intention. The explanation of the link between intending to do something and doing something intentionally, or with a certain intention, is especially critical for apprehending this proposal. I claim is that to argue in favor of a philosophically interesting theory of intention involves subscribing the suggestion of analyzing the content of an intention in terms of the notion of an action plan. In the first two sections I discuss the theory of intention. After this, I consider two problems related to the idea of intention –elucidated from the phenomenon of both plans and planning activity. Then I sketch the basic structure of intentional action emphasizing the idea of an action plan. Finally, I discuss a philosophical problem against the claim that an intention oriented to the future constitutes a central case in the analysis of intention or acting intentionally, that is, the intention in action case or the spontaneous intentional action case.

Keywords: intentional action – intention – plans – practical reasoning.

¹ IIF/SADAF-CONICET

✉ diego.lawler@gmail.com

Lawler, Diego; (2016). Actuar intencionalmente: esbozo de una teoría de la intención. *Epistemología e Historia de la Ciencia*; 1(1); 53-86. ISSN: 2525-1198



1. El punto de partida y los problemas

¿Qué supone actuar intencionalmente? Existen diferentes maneras en las que una acción es intencional y, a veces, esas diferentes maneras presuponen distintas habilidades de parte de los agentes. Por otro lado, hay que decir que el concepto cotidiano de intención no es unívoco.² No obstante, actuar intencionalmente supone, en términos generales, actuar según una intención. Es decir, cuando un individuo actúa intencionalmente hay algo que ese individuo intenta hacer; o dicho de otra manera, existe un fuerte lazo entre intentar hacer algo y hacer ese algo intencionalmente o de modo intencional.³ La intuición encapsulada en esta presentación podría formularse así: si un individuo P realiza intencionalmente una acción (A),⁴ entonces hay algo que P intenta hacer, y lo que vuelve intencional a la realización de esa acción es precisamente la relación que mantiene el hacer de P (la realización de la acción A) con una intención relevante.⁵ Por consiguiente, dado que es esa relación de la acción con una intención

² Para advertir este hecho basta recoger, por ejemplo, el señalamiento que realiza Bennett (1966). De acuerdo con este autor, el adverbio ‘intencionalmente’ puede funcionar como una elipsis excluyente. Así, decir que una acción fue realizada intencionalmente o de acuerdo con una intención significaría decir que no fue realizada accidentalmente, involuntariamente o erróneamente. En este caso, la falta de univocidad de lo que es una intención vendría dada por el hecho de que llamar ‘intencional’ a una acción dependería estrechamente del contexto; por consiguiente, constituiría un error buscar una explicación común para todos aquellos casos en que un agente actúa intencionalmente. Sin embargo, mi impresión, como se verá en lo que sigue, es que Bennett subestima una posible caracterización de la forma básica de la intención y la acción intencional que no esté exclusivamente supeditada al contexto.

³ Esto está recogido por nuestros usos ordinarios. Actuar intencionalmente significa actuar de modo intencional, lo cual quiere decir actuar de acuerdo con una intención: según la determinación de un fin (véase Diccionario de la Lengua Española de la Real Academia Española, 1992). Por otra parte, como señala Austin (1975, p. 258) al reflexionar sobre las “estelas etimológicas”, la metáfora de ‘intención’ es la de “tender o esforzarse hacia”; de allí que actuar de acuerdo con una intención implique *realizar o llevar a cabo* esa intención, o dicho en otras palabras, *realizar* el esfuerzo guiado por la determinación de un fin.

⁴ De aquí en adelante hay que leer la letra mayúscula ‘A’ como una variable que representa diferentes acciones.

⁵ Algunos filósofos, entre otros McCann (1986, 1989, 1991) y García (1990), arguyen que un agente realiza intencionalmente A sólo cuando intenta realizar A. Este punto de vista, se denomina “*the Simple View*” (Adams, 1986; Bratman, 1987). Sin embargo, adviértase que la formulación de la intuición mencionada no requiere (o demanda) que si P hace intencionalmente A, entonces P intenta A. Los filósofos contrarios a “*the Simple View*” consideran que si bien se requiere una intención, ésta no necesita ser una intención de realizar precisamente A –siguiendo a Bratman (1987, p. 112) puede denominarse este punto de vista como “*the Single Phenomenon View*”. La versión de Bratman (1987, especialmente pp. 123-124) se basa, por ejemplo, en aflojar el vínculo entre lo que se intenta y lo que se hace intencionalmente. Más allá de esta versión, las consideraciones de quienes se oponen vienen generalmente motivadas por ejemplos del siguiente estilo: imaginemos que estoy tratando de marcar un gol de tiro libre directo, esto es, enviando el balón sobre la barrera que forman los jugadores del equipo contrario, e imaginemos que creo que fallaré el disparo. Si acierto, parecería que he lanzado intencionalmente el balón sobre la barrera, pero dada mi creencia de que no resolvería exitosamente la situación, se podría pensar que no intenté hacer precisamente eso: marcar un gol de tiro libre directo enviando el balón sobre la barrera contraria. Este tipo de ejemplos ataca la conexión entre intención y confianza, respaldada por nuestro uso ordinario del término ‘intención’ y supuesta en “*the Simple View*”: un agente que no tiene confianza en que hará A –por ejemplo, alguien que piensa que sus probabilidades de enviar el balón sobre la barrera son prácticamente nulas– puede esperar, pero no intentar, hacer A (Mele, 1989, especialmente p. 19-30). En cambio, otros filósofos, por ejemplo Audi (1991, 1993), se inclinan por una variante de esta interpretación, encargándose de recalcar que no se trata del lazo entre la intención y la confianza, sino del lazo entre la intención y cierto grado de expectativas, puesto que no todo intentar, según ellos, supone la confianza. En cualquier caso, aquí interesa particularmente señalar que todos (tanto los que suscriben como los que rechazan “*the Simple View*”)

relevante lo que de alguna manera la vuelve intencional, una cuestión prioritaria será explicar qué se entiende por la noción de intención.

Las intenciones caracterizan tanto las acciones como la mente de un agente. Así, por ejemplo, decimos que podríamos, por una parte, realizar intencionalmente cierta acción (A) o realizar esa acción (A) con cierta intención y, por otra, tener por la mañana la intención de realizar cierta acción (A) por la tarde.⁶ Este retrato indica de manera amplia cuál es la tarea de cualquier teoría de la intención: considerar y explicar cómo se relacionan las intenciones qua estados mentales con las acciones realizadas intencionalmente o con cierta intención (Bratman, 1984, 1987, 2007). En consecuencia, qué es una intención dependerá de alguna manera de cómo se resuelva esta tarea, o dicho en los términos con que se abriría esta sección, dependerá de cómo se dé cuenta del lazo que existe entre intentar hacer algo y hacer ese algo intencionalmente o con una cierta intención.

Existe una tradición filosófica que desarrolla suficientemente esta labor. Se trata de una tradición que elabora una idea de intención que tiene sus raíces en Aristóteles (1970b) y sus seguidores medievales, por ejemplo, Tomás de Aquino (1988). Para estos filósofos, tener una intención es, en términos generales, tener una intención de hacer algo. En sus reflexiones sobre el razonamiento práctico, conjugan la idea de intención con una elección para actuar siguiendo un plan en un determinado tiempo y espacio.⁷ De este modo, intentar algo es elegir hacer ese algo, estar dispuesto a hacerlo y hacerlo en cuanto llegue la oportunidad, puesto que las intenciones no pueden ser demoradas.⁸ Desde esta

acordarían en que allí donde una intención de hacer A conduce de manera adecuada o correcta a realizar A, la realización de A es intencional.

⁶ Recuérdese el famoso ejemplo de Anscombe (1963) del hombre que mueve rítmicamente su brazo de arriba abajo, activando de ese modo una bomba de agua, con lo que se llena con agua envenenada un tanque de agua de una casa con el fin de eliminar a sus habitantes. En este caso hablamos de la realización intencional de esa acción (bombear el agua envenenada) o de la realización de esa acción con una cierta intención (envenenar a los habitantes de la casa). Sin embargo, también podemos emplear la noción de intención para caracterizar la mente de un agente. Imagínese la siguiente situación: sentado delante de mi ordenador formo la intención a las 15 horas de prepararme un té a las 17 horas. Cuando el reloj me indica que son las 17 horas, mi decisión o resolución anterior me conduce a dejar la computadora, trasladarme hasta la cocina y preparar el té según el procedimiento que habitualmente sigo para hacer un buen té. En este caso, la intención caracteriza mi mente. En estos dos ejemplos la acción de un agente está precedida y determinada por una decisión o resolución. Es habitual denominar a estas intenciones que preceden y causan las conductas de los agentes con el nombre de 'intenciones previas'. Y algunos autores, por ejemplo Searle (1983, especialmente pp. 79-111) y Brandom (1994, particularmente p. 244), las contraponen a las denominadas 'intenciones en la acción'. Un caso de acción intencional causada por una intención en la acción sería el siguiente: me detiene una ciudadana extranjera para preguntarme dónde está la estación de trenes de Retiro y yo respondo sin pensar. Aunque mi actuación no está precedida por una intención previa, intento, sin embargo, responder la pregunta y doy la dirección que me es solicitada. Más adelante veremos que esta idea de intención en la acción es algo problemática.

⁷ Véase, por ejemplo, Aristóteles (1970b, 1139 *a* y *b*, pp. 89-90). La idea aristotélica, luego repensada por Tomás de Aquino (1988), puede resumirse en el siguiente esquema: (a) la elección es un elemento esencial en la acción humana. En cierto sentido, se puede explicar la acción a través de explicar la elección. Ahora bien, (b) las elecciones se explican dando cuenta de la formación de las intenciones en los procesos de deliberación. Y (c) estos procesos suponen la formación de planes de acción. Por consiguiente, (d) las intenciones suponen la formación de planes que explican las elecciones que dan cuenta de las acciones. Para una discusión exhaustiva reciente de las fuentes propiamente aristotélicas, véase Ringbom (2002) y su propuesta de un modelo formal de cuño aristotélico para dar cuenta de las elecciones.

⁸ Austin (1975), señala precisamente este punto al analizar desde la gramática y la filología las numerosas y destacadas diferencias que muestran los términos 'deliberar', 'propósito' e 'intentar'. Así, la expresión

perspectiva, la acción intencional, esto es, la acción realizada según una intención, está esencialmente relacionada con el futuro; en cierto sentido, es una acción guiada por un objetivo (o meta) ausente, a saber, todavía no realizado, simplemente imaginado y representado. Dentro de esta tradición de pensamiento, la intención supone la articulación y resolución de un proceso de deliberación sobre cursos posibles de acción; quien forma una intención de algún modo elige sobre posibles no realizados.⁹

Según mi punto de vista, este marco interpretativo es el que ofrece mayores posibilidades para entender qué es una intención en sentido pleno del término. La intuición básica, expresada a la Austin (1975), es la siguiente: la intención implica en un sentido fuerte la planificación puesto que tener una intención equivale a tener un plan de acción –pace Searle (1983), entre otros.¹⁰ Desde esta perspectiva, la intención de realizar cierta acción (A) supone la planificación de cómo producir esa acción (A), el compromiso efectivo con los pasos preliminares y, finalmente, conduce a la realización de (A). O dicho de otro modo, los roles funcionales básicos de la intención están en estrecha conexión con la planificación. Por consiguiente, el contenido de la intención de un agente está de alguna manera constreñido por sus planes para realizar sus fines. O dicho de otro modo, lo que se intenta en una acción depende de lo que un agente planifica. Esta intuición es especialmente importante porque justifica, en cierto sentido, la transformación del análisis de la intención en el estudio de la formación y estructura de los planes de acción, sentando las bases para una interpretación adecuada (i.e. inmejorable) del contenido de la intención en el contexto de las acciones intencionales ordinarias.

En este trabajo me propongo lo siguiente, argumentar a favor de la explicación que sugiere que el análisis de la acción intencional conlleva analizar la intención como el producto de un proceso de deliberación y resolución sobre cursos posibles de acción, donde la formación misma de la intención supone un compromiso, en el futuro, con uno de esos cursos de acción. O dicho de otro modo, la noción de intención sobre el futuro es el eje de una teoría filosóficamente relevante de la intención. Mi argumentación a favor de esta afirmación supondrá reelaborar intuiciones de Bratman (1987) y Austin (1975), entre otros. Mi contribución consistirá en discutir las implicaciones que se siguen de tales intuiciones reelaboradas, así como en matizar las mismas e introducir diferencias, cuando esto sea necesario, y dar cuenta de otros aspectos que no fueron tratados en el

“Estoy deliberando’ podría usarse sólo para describir un proceso que está sucediendo” (las deliberaciones pueden ser demoradas); pero las expresiones “Estoy pretendiendo’ y ‘Me estoy proponiendo’ no podrían usarse para describir un proceso” (las intenciones y los propósitos no pueden ser demorados) (p. 255).

⁹ Esta tradición, para emplear términos que ahora son de uso común dentro de la filosofía, supone que el contenido de las actitudes proposicionales (deseos, creencias, intenciones, etc.) explica causalmente las acciones intencionales. No obstante, esta tradición no ha permanecido inmóvil; por el contrario, viene siendo enriquecida por continuas reelaboraciones. Sin embargo, desde que los trabajos de Davidson (1963, 1967, 1970) y Goldman (1970), para mencionar sólo dos autores, recuperaran la herencia aristotélica después de las críticas de Ryle (1949) y otros behavioristas a esta tradición, las intenciones han sido destacadas del resto de las actitudes proposicionales. Gracias a los trabajos de Audi (1993); Austin (1975), Brand (1984), Bratman (1987), Donagan (1987), Mele (1992), O’Shaughnessy (1980), Searle (1983), y Velleman (1989), entre otros, las intenciones han adquirido no sólo un estatuto ontológico propio sino también un papel central en la articulación y explicación de la acción intencional. Según mi opinión, esta rama de la tradición es la que mejor se adapta al análisis de la acción intencional. De allí que configure el trasfondo filosófico de este estudio.

¹⁰ Más adelante se discute la objeción de Searle (1983) basada en la idea de que hay acciones intencionales espontáneas que no suponen planificación, o dicho de otro modo, hay intenciones en la acción, esto es, intenciones que no conllevan plan alguno.

contexto de estas intuiciones, por ejemplo, el seguimiento de planes por referencia a la instrucción y el seguimiento de reglas, las relaciones entre las intenciones y el razonamiento práctico, entre otros. Estas reelaboraciones conducirán a dar cuenta de la estructura básica de la acción intencional en función de la noción de plan de acción.

En lo que sigue desarrollaré el andamio conceptual que sostiene la intuición básica de esta tradición. Abordaré, entonces, la cuestión de qué es una intención en relación con la tarea de cualquier teoría de la intención, a saber, dar cuenta de ella en tanto que actitud proposicional que caracteriza tanto la mente como la acción. El propósito principal de esta tarea es colocar las vigas conceptuales necesarias para caracterizar adecuadamente el contenido de las intenciones en el caso de las acciones humanas ordinarias. Una viga importante de este andamio consistirá en ofrecer una perspectiva filosófica que vuelva plausibles afirmaciones del siguiente estilo: intentar implica tener un plan, y su conversa, tener un plan implica, al menos, tener una intención.¹¹ Sin embargo, la discusión que sigue no tiene la pretensión de ser una discusión exhaustiva sobre las distintas formas que tenemos de referirnos a qué es una intención y de llamar a una acción ‘acción intencional’; por el contrario, solamente procura dar cuenta de las nociones de intención y acción intencional desde un punto de vista filosófico que resulte suficiente para esclarecer la estructura básica de las acciones ordinarias. Estructura que se refleja grosso modo en la siguiente afirmación: actuar intencionalmente supone actuar con la idea de hacer algo para realizar otra cosa –esto es, actuar de acuerdo con una intención cuyo contenido está articulado según esta modalidad instrumental. Por consiguiente, mi análisis estará articulado alrededor de esta clase de acciones.

Cualquier teoría de la intención se enfrenta al siguiente desafío: explicar cómo se relacionan las intenciones en su condición de estados mentales con la caracterización de las acciones realizadas intencionalmente o con una cierta intención. En nuestra vida cotidiana es habitual que empleemos el término ‘intención’ de acuerdo con este doble papel adjetivo (i.e. como caracterización de la mente y de la acción) porque nos provee una base fiable para predecir la conducta de los otros y coordinar nuestros planes y proyectos con los suyos. Esto nos transmite la sugerencia de que esta doble función está de alguna manera asegurada por nuestra concepción de sentido común sobre la intención. Cuando se sigue esta concepción, la noción de intención se muestra estrechamente enlazada con el fenómeno de los planes y la planificación (Bratman, 1987, p. 2). Desde esta perspectiva, el esbozo de una teoría de la intención no podría desentenderse de la relación de estos fenómenos con la intención. Sin embargo, ¿de qué modo nuestro sentido común vincula el fenómeno de los planes y la planificación con la noción de intención? Una pregunta que en cierto sentido no es independiente de esta otra, ¿qué papel desempeñan los planes en nuestras vidas para ocupar un lugar tan importante en cualquier teoría de la intención? Veamos estas dos cuestiones.

Nuestro sentido común nos indica que somos agentes planificadores: nos proponemos objetivos y trazamos planes para alcanzarlos que guían nuestra conducta posterior. Nuestra relación con el futuro está mediada por la representación y

¹¹ La primera aseveración está relacionada con el hecho de que los agentes racionales planifican sobre la base de lo que intentan –algunos filósofos, por ejemplo Audi (1991, 1993) y García (1990, 1991), agregarían: así como sobre la base de lo que creen. La segunda está vinculada con la idea general de que adoptar un plan supone adoptar una serie de intenciones; por tanto, actuar siguiendo un plan es actuar con ciertas intenciones (por supuesto, no todo lo que figura en los planes es algo que se intenta). Más adelante preciso cómo entender la noción de plan involucrada en esta segunda idea.

actualización de planes simples o complejos, que construimos con el propósito de coordinarnos intra e interpersonalmente a lo largo del tiempo. De ahí que sin planes ni planificación nuestras vidas no puedan extenderse más allá del presente. Gracias a nuestras capacidades de formar y ejecutar planes y de actuar con un propósito, estamos en condiciones de moldear nuestra conducta de acuerdo con la reflexión racional y la deliberación a partir de recursos limitados (Bratman, 1987, especialmente p. 3; 2000, especialmente pp. 35-61).

Entonces, qué respuesta puede darse a la pregunta: ¿Por qué una teoría de la intención no es independiente de una teoría de los planes y la planificación? Para decirlo de manera directa: porque formar una intención sobre el futuro es en cierto sentido elaborar un plan de acción. Esta proposición presupone que la clave para explicar la relación que mantienen las intenciones qua estados mentales con las acciones realizadas intencionalmente o con cierta intención reside en elucidar la noción de intención sobre el futuro, esto es, la idea de intención qua estado mental. De esto se sigue que una teoría de la intención es ante todo una teoría de la intención sobre el futuro.¹² En consecuencia,

¹² La idea de intención sobre el futuro es una idea que conlleva una revisión del modelo conceptual más extendido en teoría de la acción. De hecho, Bratman (1987, pp. 4-49) invierte una buena parte de su estudio en hacer plausible la dimensión *planificadora* de la intención. Ahora bien, a pesar de la obvia importancia que tienen las intenciones sobre el futuro en nuestra vida cotidiana, ¿por qué se necesita un argumento a su favor? Simplificando, podemos reconstruir el problema principal de la siguiente manera: (a) las intenciones sobre el futuro suponen un compromiso incorporado en la actividad intencional presente, a saber, si formo ahora una intención sobre el futuro, mi actividad presente se adecuará a ciertos estreñimientos; no obstante, (b) dicha actividad será compatible con un rango enorme de posibles conductas futuras; por tanto, sólo un rango pequeño de ellas pertenecerá al ámbito de lo que ahora intento; entonces, dado (a) y (b), tenemos (c): mi compromiso con la acción futura excederá el ámbito de mi actividad intencional presente. El corolario es que este tipo de compromiso podría ser metafísicamente censurable (porque involucraría alguna forma de acción a distancia) o racionalmente inaceptable (puesto que supondría la imposibilidad de revocar la intención). Sin embargo, como bien apunta Bratman (1987, p. 6), este problema pierde peso si se ataca el marco general que vuelve dudosa la idea de intención sobre el futuro. Se trata de un marco general que suscribe la teoría de la intención basada en deseos y creencias; supone que la acción intencional o la acción realizada con una determinada intención se explica en función de la relación que mantiene esa acción con las creencias y deseos del agente –para algunos, como Davidson (1980) y Goldman (1970), esa relación es una relación causal, y para otros, como Anscombe (1963), no lo es. Sin embargo, aquí importa que este modelo no admite con estatus ontológico propio lo que nuestro esquema psicológico ordinario nos muestra como obvio: el estado mental de intentar actuar, esto es, la intención sobre el futuro. Por el contrario, este modelo es profundamente escéptico con respecto a la existencia de esas intenciones; las intenciones sobre el futuro se reducirían a creencias y deseos apropiados, o mejor, a un cierto complejo de deseos y creencias. De esto se sigue que sólo hay que tener en cuenta la intención tal como aparece en la acción. Sin embargo, de acuerdo con Bratman (1987, p. 8), las intenciones sobre el futuro pueden defenderse en detrimento de este modelo si consideramos seriamente la idea de que somos agentes que hacemos planes; una idea que el modelo reduccionista descrito está incapacitado para albergar. El argumento de Bratman (1987) está trazado según la siguiente estrategia: (a) mostrar que las intenciones sobre el futuro están incorporadas en una red de disposiciones y regularidades diferentes de las que involucran deseos y creencias; si (a) puede mostrarse, entonces se sigue (b): las intenciones tendrían el estatus de estados mentales distintivos, a saber, un estatus análogo al de las creencias y los deseos. Esta estrategia supone dos resultados. Por una parte, en contraposición a la dimensión descriptiva del modelo conceptual extendido, supone afirmar que las intenciones en general y las intenciones sobre el futuro en particular son estados distintivos de la mente. Por otra parte, en contraposición a la dimensión normativa del modelo conceptual extendido, supone aseverar que las intenciones previas están en condiciones de proveer razones racionalmente relevantes para las intenciones derivadas y las acciones. Esto último acarrea un cambio importante en la concepción de la racionalidad práctica (Bratman, 1987, p. 23). En este trabajo se suscribe en sentido amplio esta línea. Por consiguiente, no se argumenta a favor de ella en el contexto de su debate con el modelo conceptual más tradicional.

una vez desplegado el contenido de esta noción se podrá dar cuenta de su articulación con el concepto de acción intencional o acción realizada según una intención.

2. La noción de intención sobre el futuro como eje de una teoría de la intención

La idea de que una teoría de la intención es básicamente una teoría de la intención sobre el futuro está sostenida por una estructura argumentativa cuya reconstrucción podría realizarse en los siguientes términos. Dado que somos agentes que hacemos planes, las intenciones sobre el futuro representan casos centrales de intención en nuestro esquema psicológico de sentido común (ordinario). O dicho de otro modo, porque somos seres que hacemos planes, la idea de intención sobre el futuro no es una idea problemática. Desde esta perspectiva, una teoría de la intención no es independiente de una teoría de los planes y la planificación porque es posible dar cuenta, al menos en sentido mínimo, del actuar intencional o del actuar con una cierta intención recurriendo a la idea de actuar siguiendo un plan o curso más o menos estructurado de acción. Es decir, el desarrollo completo de una teoría de la intención no sería posible sin la planificación, los planes y los roles que ellos desempeñan en nuestras prácticas.

Sin embargo, ¿cuál es la relación entre las intenciones sobre el futuro y los planes? ¿Cómo dar cuenta de la ligazón entre la noción de intención sobre el futuro y la de plan? Si esta relación no se aclara, se podría argüir, entre otras cosas, que este enfoque está viciado de circularidad. Brevemente expuesta, la impugnación por circularidad supone que se comete un error conceptual si se pretende explicar la noción de intención sobre el futuro a través de la noción de plan y ésta última por medio de la primera. Este error repercutiría además en la elucidación del concepto de acción intencional, puesto que éste se define por medio de esas nociones. No obstante, esta objeción se despejaría si se respondiera adecuadamente la pregunta arriba formulada. Si bien no hay una única forma acertada de contestarla, mi impresión es que la siguiente aseveración de Bratman (1987, p. 8) encapsula la explicación más atractiva para los objetivos de este trabajo: la relación (e interdependencia) entre el fenómeno de los planes y la planificación, por una parte, y la noción de intención sobre el futuro, por otra, se entiende en función de que las intenciones sobre el futuro son “los bloques de construcción de esos planes; y los planes son obviamente intenciones”; y aún más, “es poco probable que hagamos justicia al fenómeno central de los planes y la planificación, si no reconocemos esa actitud distintiva (la actitud distintiva de intentar realizar una acción o intentar actuar)”. Pero, ¿qué significado tienen estas aseveraciones? ¿Cómo pueden interpretarse desde un enfoque constructivo llano?

Una forma fructífera y rápida de capturar el enlace entre intenciones y planes consiste en prestar atención a la psicología del sentido común incorporada en nuestro

Argumentos, en mi modesto juicio definitivos, a favor de las intenciones sobre el futuro en la caracterización de la acción intencional pueden consultarse en Austin (1975), Alston (1986), Brand (1984), Bratman (1987), Mele (1992) y Mele y Moser (1994), entre otros. Más cerca en el tiempo, Bumpus (2001, especialmente p. 190-96) ha sugerido una tercera posición: las intenciones son estados distintivos pero no desempeñan un rol crucial en la explicación de la acción intencional. Sin embargo, la autora reconoce que por ahora se trata de una alternativa muy poco explorada y, por tanto, todavía inmadura para funcionar como objeción de peso.

lenguaje ordinario. Imaginemos una situación cotidiana como la siguiente: estamos en casa y vemos a nuestra pareja dirigirse al cobertizo, tomar la escalera y la tijera de podar y encaminarse hacia el sector del jardín donde están los frutales. Nuestra pareja pretende hacer algo: tiene una intención sobre el futuro (en este caso, se trataría del futuro próximo). Desde su punto de vista tiene obviamente una idea más o menos clara de qué está haciendo, o dicho de un modo coloquial, sabe qué lleva entre manos.¹³ En cierto sentido, hay que suponer que tiene un plan, esto es, algo así como una representación de una secuencia de acciones que está tratando de llevar a cabo. Según Austin (1975, p. 258), precisamente “cuando llamamos la atención sobre este aspecto de la acción, usamos las palabras conectadas con la intención”. Pero, ¿qué quiere decir Austin con este comentario? ¿Cuál es su sugerencia? Su sugerencia es que en general nuestro uso ordinario de gran parte de los verbos activos, como de las palabras vinculadas al campo semántico del término ‘intención’, incluye o tiene alguna relación con la idea de un plan o proyecto que debe realizarse o llevarse a cabo. Dicho sintéticamente, el término ‘intención’ siempre pone en juego la noción de que hay una cierta idea o representación de lo que se está haciendo. Austin lo expresa con la siguiente metáfora: tenemos una idea de lo que estamos haciendo “como si fuera una lámpara de minero en nuestra frente que ilumina siempre tanto delante como avanzamos” (1975, p. 259).

Esta metáfora elucida un aspecto que no hay que marginar cuando se reflexiona sobre la idea de un plan de operaciones implicado genéricamente en la noción de intención. Si la idea de lo que estamos haciendo es como una “lámpara de minero”, entonces su iluminación estará acotada. Pero, ¿qué significa esto para nuestra noción sobre lo que estamos haciendo? Significa que nuestra idea de lo que estamos haciendo tiene por regla general una extensión y precisión variable, o dicho con otras palabras, ella es siempre una idea en cierto sentido parcial. Veamos cómo explota el mismo Austin su metáfora:

La única regla general es que la iluminación es siempre limitada, y ello de diversas formas. Nunca se extenderá indefinidamente adelante. Naturalmente, no todo lo que se ha de seguir o hacerse en adelante es lo que yo estoy pretendiendo hacer, sino quizá consecuencias o resultados de ello. Además, no ilumina todos mis alrededores. Sea lo que fuere lo que estoy haciendo se está y debe estar haciéndose en medio de un trasfondo de circunstancias (incluyendo naturalmente actividades de otros agentes). Esto es lo que necesita cuidado, para guardarnos de tropiezos, trastornos, accidentes (1975, p. 259).

Las enseñanzas de Austin nos dicen que las intenciones suponen una idea más o menos clara de lo que se está haciendo y que esa “lámpara de minero” tiene un poder

¹³ Recordemos que, como apuntó Anscombe (1963), el hecho de que nuestra pareja sepa qué está haciendo, o tenga alguna idea de qué está haciendo, no es algo que sepa como resultado de la observación directa o indirecta de su conducta; salvo, claro está, en los casos extremos, en los cuales, para decirlo con palabras de Austin (1975, p. 258), “se da el caso de que *descubro* lo que he hecho o *llego a darme cuenta* de que estoy o he estado haciéndolo de esa forma”. Por consiguiente, “[mientras] voy por la vida haciendo, como suponemos, una cosa tras otra, en general siempre tengo una idea –alguna idea, mi idea, o imagen, o noción, o concepción– de en qué estoy, en qué me ocupo, qué llevo entre manos, o en general, ‘qué estoy haciendo’” (Austin, 1975, p. 258).

limitado: las intenciones suponen un plan parcial.¹⁴ Es decir, si no hay nada que intentas hacer, entonces no hay ningún plan parcial que hayas formado y planees realizar; por el contrario, si has formado un plan y planeas llevarlo a cabo, entonces intentas algo: te conduces según tu “lámpara de minero”. Por supuesto, siempre actúas en medio de un trasfondo de circunstancias: el mundo puede cambiar de forma no prevista; por tanto, tu “lámpara de minero” recorta su propia zona de oscuridad, esto es, entre tu idea de lo que llevas entre manos y su realización, está el boicot de las contingencias del mundo. Si tu idea de lo que llevas entre manos no es lo suficientemente parcial y flexible, resultará difícil que tu “lámpara de minero” guíe sin tropiezos tu camino.¹⁵ Una vez que se ha asegurado este punto, se está en condiciones de volver al análisis del enlace entre intenciones y planes.

Más arriba citaba a Bratman (1987, p. 8) para suscribir la idea de que las intenciones sobre el futuro son los bloques constitutivos de nuestros planes y que éstos son de una forma ostensible o acentuada intenciones. Pero antes de aclarar el significado de esta idea, es menester interrogarse sobre un asunto que las propias enseñanzas de Austin han puesto de manifiesto: ¿qué noción de plan está en juego? ¿Se trata de la noción de plan parcial que nos propone Austin? A continuación realizaré un conjunto de distinciones concernientes a la noción de plan con el propósito de disponer el escenario para una interpretación adecuada del vínculo entre intenciones y planes.

En primer lugar, conviene señalar que hay diferentes nociones teóricas relacionadas con la idea de plan. Para nuestros propósitos interesa distinguir las siguientes. Por una parte, un plan es una estructura abstracta que puede figurarse a través de distintos medios representacionales; así, hablamos de un plan para armar una bicicleta, o un plan para construir una casa, expresados, por ejemplo, narrativa o pictóricamente. Este sentido vincula la idea de plan con un esquema abstracto que entraña relaciones instrumentales entre tipos de acciones y ciertos objetivos. Estos planes pueden compartirse entre diferentes agentes y pueden llevarse a cabo por uno o varios de ellos. Bratman (1987, p 29) asocia esta idea de plan con la de una cierta clase de receta o prescripción: tengo un plan para desarmar carburadores, esto es, dispongo de un procedimiento adecuado para hacer eso o conseguir ese fin.¹⁶ Desde esta perspectiva, disponer de planes no supone necesariamente llevarlos a cabo. Uno puede tener un plan para A y no planear hacer A (Brand, 1984); por ejemplo, en este momento almaceno en mi memoria un plan para construir un estanque, cuya representación pictórica y narrativa guardo en el cajón inferior de mi escritorio, pero no planifico hacer ahora un estanque.¹⁷ La intuición subyacente es que hay un contraste entre la mera posesión vis à

¹⁴ Otra idea interesante de Austin, que rubrica el enlace entre intenciones y planes, reside en su indicación sobre el “efecto conexivo” que entraña el término ‘intención’. “[C]uando el que mete la mano en el caja del dinero afirma que *pretendió todo el tiempo* devolver el dinero, lo que está afirmando es que su acción –la acción en la que estaba ocupado– ha de juzgarse *como un todo*, no precisamente una parte de ella seccionada del todo” (Austin, 1975, pp. 259-260). Dicho en nuestro lenguaje, lo que afirma quien así procede es que actuaba según un plan.

¹⁵ La metáfora de la “lámpara de minero” no debe dar la impresión de que sólo son parciales los planes, para decirlo de alguna manera, que formamos para el día o la semana; por el contrario, también son parciales aquellos planes que estructuran gran parte de nuestras vidas y que generalmente llamamos ‘proyectos’: formar una familia, perseguir una carrera en el mundo del arte, etc. Estos planes, para decirlo coloquialmente, tienen que ir rellenándose a medida que pasa el tiempo.

¹⁶ Véase también Audi (1991, 1993), García (1991) y Mele y Moser (1994), entre otros.

¹⁷ En esta interpretación, disponer de un plan supone en cierto sentido conocerlo. Cuando decimos “Tengo un plan para construir barriletes”, implicamos habitualmente la siguiente expresión “Conozco un plan para

vis el compromiso motivacional con un plan. La mera posesión no implica motivación. La idea de tener un plan como mera posesión no juega la misma clase de papel en la explicación de la conducta intencional que la idea de tener un plan que supone motivación.

Por otra parte, un plan es cierta clase de estado mental (Bratman, 1987, p. 28). En este sentido, tener un plan para hacer A supone que tiene lugar la realización psicológica de la estructura abstracta implicada en el hacer A. Los planes *qua* estados mentales son los que generan la clase de adecuada de compromiso con la acción. Desde esta perspectiva, sólo “tengo un plan para hacer A si en mí caso es verdad que planeo hacer A” (Bratman, 1987, p. 29). A diferencia de la noción anterior, esta idea de tener un plan conlleva, además del contenido representacional sobre el curso de acción, el componente motivacional apropiado para realizar ese curso de acción planeado. Los planes *qua* estados mentales muestran propiedades propias de las intenciones: poseen cierta inercia por medio de la cual resisten la reconsideración, controlan la conducta y funcionan como insumos para razonamientos prácticos posteriores y la planificación de otros planes.¹⁸

Esta segunda idea de qué es un plan es la que resulta más atractiva para el análisis de la intención y la acción intencional. En consecuencia, conviene leer la sugerencia técnica de Bratman (1987, p. 8) de que los planes son de manera ostensible intenciones teniendo en cuenta estas observaciones. Pero, ¿qué significa que las intenciones sobre el futuro son los bloques constitutivos básicos de nuestros planes? De alguna manera significa decir que adoptar un plan supone adoptar una intención sobre el futuro y actuar siguiendo un plan, actuar con ciertas intenciones. De este modo las dos afirmaciones siguientes se asegurarían mutuamente: intentar implica tener un plan (de hecho, los agentes racionales planifican en cierta medida sobre la base de lo que intentan) y tener un plan implica al menos una intención sobre el futuro (claro está, no todo lo que figura en un plan es algo que se intenta). Finalmente, las enseñanzas de Austin agregarían que este plan es siempre en cierto sentido un plan parcial: tengo un plan para reparar el carburador del automóvil (un plan de reparación) que voy llenando, ajustando y modificando a medida que realizo y avanzo en la reparación del carburador. Desde este punto de vista, un plan es parcial en dos sentidos: es temporalmente parcial y estructuralmente parcial. Necesita ser temporalmente parcial porque entre la construcción y adopción del plan, por una parte, y su ejecución, por otra, el mundo puede cambiar de forma no prevista. Si no fuese temporalmente parcial, no podría ajustarse en tiempo real su realización a las contingencias del mundo. Por otro lado, un plan es estructuralmente parcial en el sentido siguiente: a menudo un agente se propone un fin y deja para más tarde la deliberación sobre los medios para ese fin; a medida que transcurre el tiempo el plan debe ser rellenado con subplanes que son tan vastos como el agente crea necesario para ejecutar con éxito el plan. La naturaleza parcial de los planes

construir barriletes; por ejemplo, lo tengo en mi memoria y lo reproduzco a voluntad”. En ninguno de los casos eso significa que sólo planifique hacerlo o llevarlo a cabo.

¹⁸ Por una parte, acabo de afirmar, acompañando hasta cierto punto a Bratman (1987), que tener un plan para hacer A supone que tiene lugar la realización psicológica de la estructura abstracta de cursos de acción implicada en el hacer A; y, por otra parte, asevero que los planes *qua* estados mentales generan la clase adecuada de compromiso con la acción. Entonces, adviértase que con la realización psicológica del plan lo que se intenta (i.e el contenido de la intención) no es la estructura abstracta, sino los cursos de acción contenidos en ella; de lo contrario, no podríamos explicar cómo tiene lugar la clase adecuada de compromiso con la acción.

entraña la organización jerárquica de su estructura, a saber, los planes acerca de metas u objetivos incorporan planes acerca de medios y pasos preliminares.^{19 20}

En resumen, hasta aquí he puesto de manifiesto que las intenciones son estados mentales que caracterizan tanto nuestra mente como nuestra conducta; además, he señalado que cualquier teoría de la intención debe tratar estos dos aspectos y mostrar su articulación. A su vez, he explicitado que, según nuestra concepción de sentido común, la intención qua estado mental de intentar actuar desempeña un papel explicativo central en la intención y la acción intencional en general, esto es, es posible dar cuenta de lo que es hacer algo intencionalmente o con una cierta intención a partir de elucidar lo que es intentar hacer algo (tener una intención sobre el futuro). La ocasión para un movimiento de esta clase venía dada porque básicamente somos agentes planificadores. Disponemos de recursos limitados y soportamos la presión de coordinarnos en el tiempo individual (intrapersonalmente) y colectivamente (interpersonalmente); puesto que esto demanda la extensión de la reflexión racional y la deliberación más allá del presente, los planes parciales sobre el futuro juegan un papel medular en nuestras vidas. De esto se sigue que si nuestra comprensión de la intención es un asunto de nuestra comprensión de la intención dirigida hacia el futuro, entonces no es independiente de los planes parciales y del papel que desempeñan en nuestras vidas. Esto último resulta, en consecuencia, una pieza clave para el desarrollo de una teoría de la intención.

3. La teoría de la intención resultante

Si nuestra comprensión de la intención es fundamentalmente un asunto de nuestra comprensión de la intención dirigida hacia el futuro, ¿qué clase de teoría de la intención resulta de esta concepción? Se trata de una teoría de la intención que puede articularse en dos dimensiones: una descriptiva y otra normativa. La dimensión descriptiva se resume en la siguiente aseveración: las intenciones en general, y las intenciones sobre el futuro en particular, son estados distintivos y verdaderos de la mente junto con las creencias y los deseos. Ahora bien, ¿qué significado tiene esta afirmación? Significa que las intenciones están incorporadas en regularidades y disposiciones subyacentes de una clase particular, esto es, diferente de las propias de los deseos y

¹⁹ Un ejemplo que recogería estas características podría ser el siguiente: imaginemos que formo hoy jueves la intención de ir a la cancha de fútbol este domingo. Por supuesto, no necesito disponer ahora de un plan completo; por el contrario, dejo para el sábado la deliberación sobre qué partido de fútbol ver, cómo obtener las entradas, a qué amigo invitar, etc. A medida que se acerca el momento voy especificando los subplanes respectivos, los medios necesarios, los pasos preliminares, el ajuste de este plan con otros que tenga en mente, etc. Para una discusión interesante de este punto, véase Bratman, Israel y Pollack (1988). La naturaleza jerárquica de los planes está bien recogida en la discusión sobre planes y estructura de la conducta de Miller, Galanter y Pribram (1983, especialmente capítulo 2).

²⁰ Estas características entrañan, además, razones pragmáticas de peso. Véase esta extensa cita de Bratman (1987, pp. 29-30): “Por un lado, necesitamos coordinar nuestras actividades, tanto al interior de nuestras propias vidas así como socialmente, a saber, con otras vidas. Además, necesitamos hacer esto de una forma que sea compatible con nuestras capacidades limitadas para la deliberación y el procesamiento de la información. Incluso más, dadas estas mismas limitaciones, necesitamos que la deliberación previa dé forma a nuestra conducta futura. Esto conforma una vía argumentativa a favor de nuestra condición de seres planificadores. Por otra parte, el mundo cambia de una manera que no estamos en condiciones de anticipar; por consiguiente, planes muy detallados sobre el futuro mediato son de muy poca utilidad -no merece que nos preocupemos por eso. Planes parciales para el futuro, estructurados y jerarquizados, constituyen nuestra solución de compromiso”.

creencias, que las conectan con otros estados mentales, procesos psicológicos y actividades y con inputs (percepciones) y outputs (acciones) apropiados.²¹ La dimensión normativa, por su parte, recoge las consecuencias que tiene para la racionalidad práctica sostener esta dimensión descriptiva. En los párrafos siguientes exploraré con detenimiento estas dos cuestiones.

¿Cuáles son las características de esa red de regularidades y disposiciones, distintas de las propias del deseo y las creencias, a la que se incorporan las intenciones? Se trata de una red de regularidades y disposiciones ligada esencialmente al papel de los planes en el razonamiento práctico y la acción. La idea general es la siguiente: tener una intención supone tener cierta clase de compromiso, algo que especialmente se pone en evidencia en el caso de las intenciones sobre el futuro.²² Este compromiso presenta un aspecto relacionado con la volición y otro vinculado a la estructura del razonamiento práctico de un agente.²³ De acuerdo con el aspecto volitivo, tener una intención no es

²¹ En general, el entendimiento que nos provee el sentido común sobre nuestros distintos tipos de estados mentales depende de la presuposición de las regularidades apropiadas para cada uno de ellos. En el caso de las intenciones, se trata de regularidades y disposiciones, asociadas con el papel de los planes parciales, que son críticas para el razonamiento práctico de agentes racionales limitados. Véase Bratman (1987, especialmente pp. 14-27).

²² El compromiso como marca de la intención está ya en Tomás de Aquino (1988, I-II, 12, 1-4, pp. 142-146).

²³ Creo que un esquemático ejemplo puede destacar rápidamente la dimensión del compromiso que supone la intención. Imaginemos la siguiente situación: Javier disfruta los efectos inmediatos del whisky escocés, especialmente del buen whisky escocés. Sin embargo, Javier no disfruta los efectos que acarrea beber demasiado whisky: no disfruta la resaca que habitualmente sufre cuando satisface excesivamente su primer deseo. Podríamos suponer una circunstancia en que Javier formara la intención de beber whisky y a su vez formara la intención de evitar contraer una resaca. Supongamos ahora que Javier ha salido esta noche, ha bebido unos cuantos whiskies y su amiga Valentina le ofrece otro, una malta de alta calidad de dieciocho años de reposo. Podríamos distinguir tres maneras en que Javier puede reaccionar ante esta invitación de su amiga. En primer lugar, Javier puede haber bebido tanto que ya ha olvidado su intención de evitar la resaca del día después; incluso puede haber olvidado que disfruta bebiendo whisky. Simplemente sigue sus inclinaciones y bebe el whisky. Su acción no es una respuesta ni a su compromiso práctico de beber whisky ni a su compromiso práctico de evitar la resaca –ambos implicados por las intenciones atribuidas. Por el contrario, Javier actúa meramente de acuerdo con sus disposiciones, de una forma puramente apetitiva. Una segunda posibilidad consistiría en que Javier declinara el whisky porque entreviera que otro vaso podría originarle definitivamente resaca. Existen diferentes maneras en que Javier podría tomar esta decisión. Quizá Javier hace una pausa, resuelve no tener resaca y, sobre la base de esta intención previa a la acción, no toma otro whisky. Esta intención previa podría conllevar el pensamiento de un principio general, ‘me desagradan las resacas’, o podría conllevar simplemente la percepción de que otro vaso conduciría a la resaca. O quizá Javier actúa de acuerdo con una intención, ya sea sobre la base de su habilidad para reconocer la situación en la que se aplica el principio general o porque la simple percepción de esa situación demanda abstinencia. En todos estos escenarios, la acción intencional de Javier es fundamentalmente diferente de la primera situación. En cada una de ellas, su acción está motivada por el reconocimiento de un compromiso práctico a través de la intención (ya sea de un compromiso en la forma de una regla general o de la percepción del mejor curso de acción posible). En estos casos la acción de Javier es intencional en un sentido en que la acción en el primer escenario no lo es: está causada por su reconocimiento de su propio compromiso práctico. La tercera posibilidad es la siguiente: Javier toma el vaso de malta añeja, se da cuenta de su intención y del respectivo compromiso de evitar una resaca, pero elige beberla sobre la base de su intención y del respectivo compromiso con las maltas escocesas añejas de excelente calidad. Esta elección podría poseer la forma de cualquiera de las descritas en el segundo escenario – i.e. podría ser articulada a lo largo de distinciones entre la aplicación de una regla y la percepción del mejor curso de acción posible. En este tercer escenario, Javier ha respondido nuevamente a una toma de conciencia de su compromiso práctico basado precisamente en una intención. O dicho de otro modo, su reconocimiento del compromiso práctico ha jugado un rol causal en su economía conductual. El asunto es que hay un sentido de ‘intencional’ de acuerdo con el cual decimos que Javier no actuó

análogo a tener un deseo; a diferencia del deseo, que es un mero influyente potencial de la conducta, la intención es una actitud mental controladora de la conducta.²⁴ Según el segundo aspecto, las intenciones poseen cierta estabilidad o inercia, esto es, existe una tendencia a mantener las intenciones y a no reconsiderarlas, excepto que se presente información novedosa y relevante. Esto último les permite funcionar como inputs para otros razonamientos prácticos, a saber, se puede razonar sobre intenciones futuras desde intenciones previas, sobre intenciones específicas a partir de intenciones generales y sobre medios o pasos preliminares a partir de un objetivo intentado. Sin esta idea del compromiso que supone la intención sobre el futuro resultaría difícil entender el papel de coordinación intrapersonal y social que cumplen los planes en el razonamiento práctico y la acción. Dicho de otro modo, dado que somos agentes limitados, planificamos: elaboramos planes parciales que ajustamos, combinamos y empleamos para modificar otros planes. Esta actividad supone una coordinación racional, coordinación que se aprehende y explica entendiendo la intención como una clase de compromiso que desempeña un rol causal distintivo. Desde este punto de vista, la intención se encontraría en condiciones de proveer cierto tipo de consideraciones (por ejemplo, razones) para la acción y formación de otras intenciones. Esta consecuencia normativa que afecta a la racionalidad práctica es el corolario de la afirmación descriptiva anterior, afirmación que rescata el papel básico de la intención en el razonamiento práctico a través de la idea de compromiso.

El contexto que acoge estas dos dimensiones (descriptiva y normativa) puede sintéticamente describirse de la siguiente manera. Las intenciones relacionadas con nuestras acciones futuras son elementos que forman parte de extensos planes o proyectos.²⁵ Dada nuestra condición de agentes limitados, dichas intenciones facilitan nuestra coordinación personal y social, ayudando en nuestras deliberaciones a dar forma

intencionalmente en el primer escenario (cuando actuó de una forma puramente apetitiva) y sí actuó intencionalmente en el segundo y en el tercero, a pesar de que las conductas resultantes en el primero y en el tercero fuesen idénticas. Y ese sentido está en gran parte dado por el compromiso implicado. De allí se sigue que tener una intención involucra cierta clase de compromiso.

²⁴ La función de la intención de controlar la conducta se puede describir a veces en términos causales, por ejemplo: “Si intento A ahora, mi intención me conducirá normalmente al menos a tratar de hacer A” (Bratman, 1987, p. 108). En cualquier caso, el aspecto volitivo que presenta el compromiso implicado por la intención se percibe más claramente cuando se lo contrasta con el deseo. En palabras de Bratman (1987, p. 16): “[M]i deseo [...] necesita todavía ser contrapesado con otros deseos en conflicto [mientras que cuando formo una intención] normalmente no necesito sopesar otra vez los pros y contras”. Sobre la distinción entre deseos e intenciones, véase también Brand (1984, especialmente pp. 121-127).

²⁵ Los planes extensos o proyectos se refieren a las metas que nos proponemos y que abarcan periodos extensos de tiempo. Tenemos un proyecto, por ejemplo, cuando planeamos especializarnos de aquí a cinco años en el movimiento de los flujos de inversiones del sector petroquímico en *Wall Street*. Posteriormente, realizamos ese proyecto concretando planes parciales; por ejemplo, decidimos estudiar diariamente el comportamiento del *portfolio* de acciones e inversiones de las firmas más importantes del sector; planeamos la realización de unos cursos específicos en el extranjero sobre esta cuestión; desarrollamos un plan para conseguir fondos que nos permitan pagar en parte esa estancia; diseñamos un programa que modela las estrategias financieras de las firmas del sector en función de otras variables económicas relevantes; estudiamos la innovación tecnológica en el sector petroquímico; realizamos consultas con geólogos sobre la posible incorporación de nuevas zonas geográficas de explotación, etc. Como rápidamente salta a la vista, los planes extensos o proyectos suponen planes parciales complejos que a su vez entrañan otros planes parciales más simples, etc. Como ya se señaló, estos planes parciales pueden ser tan extensos como el agente considere necesario para ejecutar con éxito el plan general.

a nuestras acciones. Cuando buscamos la coordinación y la acción efectiva no deliberamos en el vacío; por el contrario, fijamos por adelantado planes parciales y tendemos a reconsiderarlos cuando debemos enfrentar algún problema no previsto o disponemos de información nueva relevante. Si no se presentan impedimentos o nueva información, los planes son conservados. En cualquier caso, tanto si son mantenidos como si son reconsiderados, los planes parciales incorporados en las intenciones son insumos para otros razonamientos prácticos. De allí que la habilidad de postular planes sea una condición básica para alcanzar objetivos complejos; para decirlo con palabras de Bratman (1987, p. 28): “Esta habilidad para disponer y encastrar planes coordinados es una clase de medio universal: es de un uso significativo en la prosecución de fines de distinta naturaleza”.²⁶

En definitiva, según el marco conceptual desarrollado, la noción emergente de intención está inextricablemente ligada a los planes y la planificación. La intención es una clase de actitud proposicional especial: es una actitud ejecutiva orientada hacia planes de acción. La intención sobre el futuro se sostiene por nuestra condición de agentes planificadores. El análisis de esta condición pone en evidencia la dimensión planificadora de la intención; algo que, por otro lado, se advierte fácilmente una vez que se admite que la acción intencional involucra la ejecución de planes previamente formados y representados.²⁷ Sin embargo, los planes, además de ejecutarse, se forman, combinan y constriñen con otros planes, por lo tanto, se modifican y actualizan. Las intenciones sobre el futuro, en su condición de bloques constitutivos elementales de esos planes, también están involucradas en procesos y actividades de esa clase. Más adelante presentaré una aproximación a la dinámica de esos procesos, cuyo foco es el análisis de las intenciones sobre el futuro en su condición de insumos y resultados de razonamientos prácticos.

En los párrafos anteriores retraté la red de regularidades y disposiciones ligadas esencialmente al papel de los planes parciales a través del análisis de la idea de compromiso que supone una intención.²⁸ En esa presentación destacué tanto el aspecto conativo (o volitivo), ligado al inicio y sostenimiento motivacional de la acción, como el vinculado al rol que pueden desempeñar las intenciones en la articulación del razonamiento práctico. Para finalizar con este tratamiento de la teoría de la intención, consideraré dos cuestiones de cierta importancia. En primer lugar, este interrogante: ¿en qué sentido y cómo nuestra idea de lo que estamos haciendo guía efectivamente lo que hacemos? La consideración de esta pregunta permitirá dar una vuelta de tuerca sobre los componentes del contenido de las intenciones sobre el futuro (i.e., la estructura interna de la “lámpara de minero”). En segundo lugar, este otro, ¿qué distancia se está dispuesto a admitir entre la representación de las acciones e instrucciones propias de un plan y la realización de ese plan para interpretar que un agente actúa intencionalmente, esto es,

²⁶ Para la aplicación de esta idea a la inteligencia artificial, véase Pollack (1992, pp. 43-68).

²⁷ A diferencia de Bratman (1987), Mele (1992, p. 114) distingue entre intenciones dirigidas hacia el futuro no inmediato (“intenciones distales”) e intenciones dirigidas hacia el futuro inmediato (“intenciones proximales”). Esta distinción, sin embargo, no afecta nuestra caracterización basada en el caso genérico de intenciones sobre el futuro, puesto que tanto las “intenciones proximales” como las “intenciones distales” son actitudes ejecutivas hacia planes. “Una intención proximal de hacer A incorpora una propensión del agente a ejecutar cierto componente representacional de la intención, que he denominado un *plan*” (Mele, 1992, p. 176).

²⁸ Este compromiso contribuye a explicar, en palabras de Bratman (1987, p. 18), “por qué sobre una intención de hacer A se apoya normalmente una creencia de que el agente realizará A”, una creencia que, por otra parte, facilita la coordinación.

actúa siguiendo ese plan? Esta pregunta lleva a examinar la cuestión del seguimiento de los planes por referencia al problema filosófico de la instrucción y el seguimiento de reglas. Veamos a continuación, en dos secciones diferentes, cada una de estas preguntas.²⁹

4. La “lámpara de minero” y el guiar de lo que hacemos

La pregunta es: ¿en qué sentido y cómo nuestra idea de lo que estamos haciendo guía efectivamente lo que hacemos? Detengámonos por un momento en la caracterización del contenido de las intenciones sobre el futuro.³⁰ La intención sobre el futuro tiene un componente representacional y un componente conativo.³¹ En cualquier caso, estos componentes o eventos configuran, en términos generales, la contribución causal de la intención a la realización de la acción.³² Como se pudo apreciar en el tratamiento de los roles funcionales propios de las intenciones sobre el futuro, el aspecto conativo tiene como función motivar y sostener la realización de la acción, mientras que el aspecto representacional tiene como papel principal representar previamente el plan de acción y guiar y monitorear su realización. Se podría considerar que el componente conativo funciona como una especie de condición de posibilidad del componente representacional; es decir, sin componente conativo no se iniciaría ni sostendría la acción. Además, sin ese componente tampoco habría ninguna estrategia viable que permitiera analizar cómo la acción es guiada y monitoreada.³³ No obstante, que se requiera el componente conativo no significa que su contenido forme parte del contenido del componente representacional. En este caso, la condición de posibilidad no forma parte del contenido de aquello que hace posible. Entonces, nuestro primer interrogante puede expresarse con mayor precisión en estos términos: ¿en qué sentido y cómo el componente representacional guía y monitorea lo que estamos haciendo?

²⁹ El orden en el tratamiento de estas cuestiones tiene que ver con que debe satisfacerse el papel de guía que cumple un plan para plantear el problema de seguir un plan. Véase, entre otros, Brand (1984, especialmente pp. 240-241).

³⁰ Adviértase que hasta aquí el análisis ha girado básicamente en torno a los roles funcionales de las intenciones sobre el futuro. En los párrafos que siguen se orientará hacia el retrato del contenido de las intenciones. No obstante, se tendrán en cuenta algunos elementos del análisis anterior.

³¹ O dicho técnicamente: “intentar es un evento mental complejo que ha de ser descompuesto en rasgos cognitivos y conativos” (Brand, 1984, p. 46).

³² Recuérdese que allí donde una intención de hacer A conduce de la manera adecuada (o correcta) a realizar A, la realización de A es intencional. Los componentes del contenido están encapsulados en la intención de hacer A; la contribución causal a la realización intencional de A está implicada en la idea de que la intención de hacer A *conduce adecuada o correctamente* a realizar A. Dicho de otro modo, cuando actuamos intencionalmente somos movidos y guiados a actuar como lo hacemos en parte por una intención apropiada. La siguiente discusión se propone ofrecer una interpretación de la cuestión: en qué sentido y cómo somos guiados a actuar como lo hacemos. Como se verá a continuación, mi intuición es que eso no es completamente independiente de la cuestión de cómo somos efectivamente movidos. Para una caracterización que en cierto sentido comparte la estructura de esta presentación del contenido de una intención, véase la idea de Davidson (1980) de que una razón efectiva está compuesta de un elemento conativo y un elemento representacional que liga la ‘conación’ a la conducta apropiada.

³³ Esto significa que es realmente difícil imaginar que las representaciones cognitivas puedan por sí mismas desempeñar el papel de motivar la conducta. La representación de un plan de acción, independientemente de cuán persistente y medular resulte, es compatible con su no-ejecución. Véase, entre otros, Brand (1984), Mele (1992) y Miller, Galanter y Pribram (1983).

De acuerdo con las discusiones anteriores, una acción intencional es una acción que se actúa siguiendo un plan.³⁴ Es decir, no es suficiente que la acción sea causada por una intención; además, se requiere que sea una acción planificada, esto es, que la representación de la acción (componente del contenido cognitivo de la intención) guíe la realización de esa acción (o dicho de otra manera, que la acción se realice siguiendo un plan).³⁵ Y una idea suficientemente robusta de cómo llevar adelante una acción (por ejemplo, desarmar el carburador), para que funcione como guía efectiva de su realización intencional, será un plan de acción para realizar dicha acción.³⁶ Sin embargo, si al ejecutar una intención soy guiado por el plan que incorpora, ¿en qué sentido y cómo el plan guía la acción?

El plan, qua componente representacional de la intención, es en sí mismo un componente complejo. A los fines de su análisis, puede subdividirse en dos representaciones diferentes, aunque complementarias. Una representación sobre la compleja actividad que el agente habrá de realizar y otra que entraña patrones más o menos detallados de movimientos corporales específicos. La primera representación tiene cierta estabilidad a lo largo de la acción y se refiere al propio plan de acción del agente. La segunda, inestable, involucra patrones más o menos detallados de movimientos corporales concretos que proporcionan información para monitorear, corregir y ajustar el proceso de realización del plan (Brand, 1984, p. 153-154). El sentido y la manera en que el contenido representacional de una intención guía y monitorea la realización intencional de una acción se comprende por referencia a estas dos representaciones. Cuando el contenido representacional figura adecuadamente en la etiología de la acción, entonces el plan (considerado en su duplicidad representacional) guía, monitorea y coordina la realización de esa acción.

La presencia adecuada del contenido representacional en la etiología de la acción significa dos cosas. En primer lugar, significa que la segunda clase de representación (representación de movimientos corporales concretos) está controlando constantemente el progreso de las acciones hacia la realización del objetivo. Se trata de una actividad de monitoreo que involucra al rol causal que se basa en la entrada, almacenamiento e interpretación de la información (o desinformación) durante la ejecución continuada del plan representado. En segundo lugar, significa que la primera clase de representación (la representación del plan de acción del agente) especifica apropiadamente las condiciones para realizar intencionalmente la acción en cuestión.³⁷ Entonces, ‘figurar adecuadamente en la etiología de una acción intencional’ describe el contenido representacional de una

³⁴ Recuérdese que esta afirmación no es nueva en la literatura sobre filosofía de la acción. Entre otros, fue sugerida por Goldman (1970) y Castañeda (1975). No obstante, las mejores elaboraciones que conozco son las siguientes: Alston (1986); Audi (1993); Brand (1984); Bratman (1987) y Mele (1992).

³⁵ Una parte importante de la satisfacción del rol causal apropiado que desempeña la intención en una acción intencional se interpreta como guiar adecuadamente la realización de la acción.

³⁶ Varios autores argumentan que, dadas las características que poseen los planes, el componente representacional de la intención no puede ser una creencia, puesto que esta última no podría desempeñar la función de monitorear el seguimiento del plan: “Las creencias son accesibles a la conciencia y relativamente estables. Sin embargo, el seguimiento y la guía de la actividad motora no es en general ni conscientemente accesible ni estable”; y más adelante concluye: “esta actitud cognitiva es ella misma demasiado compleja para ser el contenido de una creencia, construida según el sentido común.” (Brand, 1984, p. 154). Véase también Mele (1992).

³⁷ Una caracterización adecuada de las intenciones no implica que las condiciones especificadas sean condiciones suficientes. Por tanto, no se quiere decir que esta clase de representación dé lugar a condiciones suficientes para realizar intencionalmente la acción en cuestión.

intención que satisface sus roles funcionales de monitorear, guiar y coordinar esa acción.³⁸

5. El seguimiento de los planes por referencia a la instrucción y seguimiento de reglas

La cuestión de esta sección queda retratada en los siguientes interrogantes. ¿Qué distancia se está dispuesto a admitir entre, por una parte, la representación de las acciones e instrucciones propias de un plan y, por otra, la realización de ese plan, para interpretar que un agente actúa intencionalmente, esto es, actúa siguiendo ese plan? O mejor, ¿cuán estrechamente debe un agente seguir un plan para decir que ese agente realiza intencionalmente las acciones pertinentes?

Dado que la intención es una condición de la acción intencional, cuando se intenta A, se intenta de una cierta manera, esto es, de acuerdo con un plan incorporado en la intención. Según lo dicho en (2.6), todo lo que se necesita asumir es que el plan implicado en la intención de hacer A especificará adecuadamente las condiciones de esa acción intencional.³⁹ Sin embargo, la función de guiar y coordinar la acción que tiene un plan plantea otra cuestión filosóficamente importante. El contenido representacional de una intención caracteriza cierta acción como acción intencional o acción realizada según una intención cuando se sigue el plan de esa acción. La cuestión filosófica en juego reside en el contraste entre seguir un plan y actuar azarosamente de acuerdo con un plan. Se trata de una cuestión que no es anodina en ningún sentido, puesto que sugiere que hay aspectos de la acción intencional que podrían situarse y comprenderse por referencia al escenario de la instrucción y el seguimiento de reglas. Seguir una instrucción supone algo más que el mero involucrarse en acciones que acuerdan con esa instrucción. O dicho de otro modo, seguimos una regla cuando nuestras conductas son apropiadamente guiadas por esa regla. Del mismo modo, si una persona no tiene ninguna idea de cómo desarmar un carburador, sólo puede desarmarlo accidentalmente, y puesto que no seguiría ningún plan, no estaríamos inclinados a decir que ha actuado intencionalmente. Entonces, la pregunta acerca de qué significa seguir un plan está emparentada con un problema filosófico más profundo: ¿qué significa seguir reglas? ¿Qué hace de una práctica una práctica reglada? ¿Qué habilidades presupone la participación en una práctica reglada?

Mi propósito en los párrafos que siguen no es proporcionar una discusión completa de estas últimas preguntas. Por el contrario, deseo realizar una exploración rápida, aunque conceptualmente estructurada, sobre la naturaleza de las prácticas

³⁸ Adviértase que se está dejando de lado la discusión de si las intenciones han de ser o no autoreferenciales. La pregunta sobre en qué sentido y cómo el componente representacional de la intención guía la acción correspondiente, permitiría discutir si el plan tiene o no que referirse a la intención de la que forma parte; es decir, si el rol de las intenciones en las condiciones de la acción intencional podría ser o no adecuadamente satisfecho sin suponer (o suponiendo) que las intenciones son autoreferenciales. Para algunos filósofos, por ejemplo: Harman (1976, 1986) y más recientemente Searle (1983) y Donagan (1987), la guía por parte de una intención, por medio de un plan que es parte de la intención o está asociado con ella, requeriría que la intención sea autoreferencial o que al menos el plan se refiriese a la intención. Para una visión opuesta, véase Brand (1984), Bratman (1987, 2007), Mele (1992), y Mele y Moser (1994), entre otros.

³⁹ Y si recordamos la aseveración de Bratman (1987, p. 29): “Tengo un plan para hacer A solamente si es verdad de mí que planeo hacer A”, al plan (componente representacional) hay que sumar el componente motivacional respectivo.

regladas en general, con la expectativa de que nos provea de herramientas conceptuales para examinar las habilidades implicadas en algunas prácticas regladas particulares: en especial, en la práctica de seguir un plan y atribuir a alguien el seguimiento de un plan. Para ello recurriré a una presentación reciente de estos problemas realizada por Brandom (1994).⁴⁰

La distinción entre estar regido por leyes de la naturaleza y seguir una regla es una distinción bien asentada en filosofía. Por ejemplo, los planetas obedecen las leyes de la naturaleza; en cambio, el jugador de tenis de mesa sigue las reglas del tenis de mesa. El tenis de mesa es una práctica normativa o reglada; por el contrario, el movimiento de los planetas no lo es. Por consiguiente, nuestra primera cuestión es ¿qué hace que algunos eventos estén gobernados por reglas y otros regidos por leyes? O, para formular la pregunta desde un punto de vista interpretativo, ¿cuáles son los criterios de acuerdo con los cuales decidimos que algunas conductas siguen una regla?

Obsérvese lo siguiente. Por una parte, la mera regularidad de una conducta no es suficiente para identificarla como conducta reglada. Esto es bastante obvio en el caso de agentes o actores individuales. Podemos construir un autómatas cuya conducta se corresponda con las reglas de juego del truco (juego de naipes) pero negar que esa máquina siga las reglas en el sentido en que podría seguirlas una persona.⁴¹ Sin embargo, ha habido un número importante de intentos por identificar la normatividad en la regularidad social, en la regularidad entre los individuos. No obstante, la regularidad, incluso entre individuos, no puede ser el único criterio para la normatividad.⁴² O dicho de otro modo, un interés meramente descriptivo por las regularidades nunca elucidará qué prácticas son normativas (en función de alguna regularidad que pueda siempre ser encontrada). La regularidad, sea social o individual, puede ser un criterio necesario para decir que una práctica está gobernada por reglas, pero nunca será un criterio suficiente. Siguiendo a Brandom (1994) denominaré ‘regularismo’ a esta posición.

Por otra parte, sigue siendo problemático trasladarse al polo opuesto, esto es, identificar las prácticas regladas como aquellas que explícitamente establecen sus reglas. En este extremo, la deriva es hacia el bien conocido problema del regreso de las reglas.⁴³

⁴⁰ Por supuesto, deseo permanecer finalmente agnóstico respecto de la posición del propio Brandom, i.e. su fenomenalismo sobre las normas, puesto que la valoración de la misma no es asunto de este trabajo. Independientemente de mi juicio sobre ella, creo que el diagnóstico que realiza Brandom tiene un valor especial para el objetivo que aquí se persigue.

⁴¹ Por supuesto, hay un sentido en que este autómatas sigue reglas. Se trata del mismo sentido en que decimos que un programa de ordenador sigue reglas. No obstante, no está el mismo sentido en juego cuando decimos que una persona sigue reglas.

⁴² Brandom realiza una crítica a este último intento: “El intento de identificar la distinción normativa entre actuaciones correctas e incorrectas con la distinción naturalista entre actuaciones regulares e irregulares, falla porque ninguna actuación es simplemente irregular (aun en relación con un conjunto específico de actuaciones que se emplearía para evaluar su co-regularidad); por lo tanto, ninguna sería considerada incorrecta por ese criterio. La apelación a la regularidad e irregularidad podría ser normativamente explicativa sólo si hubiese alguna manera de privilegiar algunas regularidades por sobre otras –para expresarlo con otras palabras, alguna manera de decir lo es la regularidad es correcta” (1994, p. 208). La etiqueta que aplica Brandom a esta clase de intentos está dada por el término ‘regularismo’ (‘regularism’).

⁴³ Véase Brandom (1994, p. 21, p. 63 y p. 74). Si bien este argumento es obviamente wittgensteiniano, fue aplicado por primera vez con excelentes resultados por Ryle (1967). Para una formulación de este argumento, véase, entre otros, Sellars (1963). Brandom remonta los orígenes de esta interpretación errónea (a saber, el regulismo) hasta Kant: “Para [Kant], así como para la mayoría de los filósofos antes de este siglo, la reglas explícitas y los principios no son simplemente una forma entre otras que lo normativo podría asumir. Las reglas son la forma propia de lo normativo. Este punto de vista, a saber, que las propiedades

Si una práctica formula explícitamente sus reglas, entonces esas reglas necesitan ser interpretadas y aplicadas. La interpretación y aplicación es en sí misma una práctica reglada, puesto que puede ser llevada a cabo correcta o incorrectamente. Si las reglas explícitas son una condición necesaria de todas las prácticas regladas, tendríamos un regreso infinito: aplicar las reglas a cualquier práctica requeriría explicitar las reglas que gobiernan cómo tales reglas han de ser aplicadas; a su turno, la aplicación de esas reglas explícitas requeriría otras reglas explícitas, etc. Siguiendo a Brandom (1994) denominé ‘regulismo’ a este enfoque.

Para evitar el círculo vicioso que se ha descrito en el párrafo anterior, se debe suponer que cada aplicación de las reglas explícitas descansa, eventualmente, en alguna práctica normativa implícita que gobierna cómo aplicar reglas. Habría, entonces, prácticas normativas cuyas reglas están implícitas. Las reglas explícitas son las reglas que están claramente establecidas.⁴⁴ Las reglas operativas en una práctica que no están explícitas, están implícitas. De acuerdo con el argumento anterior, hay prácticas cuyas reglas están implícitas. Estas prácticas pueden ser llamadas ‘prácticas implícitas’.

La fuerza lógica del argumento contra el “regulismo” parece ser doble. Por un lado, las reglas explícitas no son condiciones de posibilidad necesarias para todas las prácticas regladas. Por otro lado, cualquier práctica reglada que emplee reglas explícitas presupone para su aplicación algunas reglas implícitas. Sin embargo, no hay que malentender esta doble fuerza lógica del argumento. En primer lugar, del hecho de que las reglas explícitas no sean condiciones necesarias de todas las prácticas regladas no se sigue que no puedan ser condiciones necesarias de algunas prácticas regladas. Podría haber algunas clases de prácticas normativas que para ser identificadas como tales supusieran necesariamente el uso de reglas explícitas. Lo que el argumento contra el ‘regulismo’ probaría es que esas clases de prácticas (que suponen el uso de reglas explícitas) descansan sobre otras prácticas, a saber, prácticas implícitamente regladas. En segundo lugar, es un error identificar las reglas explícitas con las prácticas implícitas que las hacen posibles. Para seguir las reglas del truco impresas en “El libro del Truco”, debo ser capaz de leer castellano, reconocer las distintas cartas de la baraja española, sumar y restar, etc. Todas estas son prácticas regladas y, quizá, las he aprendido sin tener que recurrir a reglas explícitamente formuladas. No obstante, ellas no son las reglas del truco. Aprender a jugar al truco leyendo “El libro del Truco” no presupone conocer esas reglas implícitamente. Esto es, una práctica explícitamente reglada no presupone la presencia (o aun la posibilidad) de una práctica implícitamente gobernada por esas mismas reglas.

Una manera interesante de entender el seguimiento de planes es considerar que la conducta reglada no está, o no es en sí misma, gobernada por reglas. Es decir, no hay un hecho que haga a una conducta reglada y a otra no. Por el contrario, aquello que hace de las conductas de alguien conductas regladas es que sean tratadas como tales (como conductas regladas) por otros (intérpretes). Este giro produce un cambio en el análisis y permite enfocar mejor nuestro asunto de interés. El problema se traslada, entonces, de la cuestión de qué debe hacer una persona para seguir una regla a la cuestión de qué

de la práctica están siempre y en todo lugar concebidas como expresiones de la obligatoriedad hacia los principios subyacentes, puede ser denominado regulismo sobre las normas” (1994, p. 19-20). La última parte de esta cita menciona la etiqueta con que Brandom identifica esta posición: ‘regulismo’ (‘regulism’).

⁴⁴ Según Brandom (1994, p. xviii): “hacer algo explícito es decirlo: darle en una forma en que pueda ser puesto en juego como una razón, o como razones que son solicitadas en su favor”.

supone interpretar que alguien está siguiendo una regla.⁴⁵ En particular, si aceptamos que el seguimiento de planes implica una práctica reglada, entonces nuestra pregunta podría ser: ¿qué está supuesto cuando interpretamos que un agente sigue un plan?

Según este enfoque interpretativo, para que tengan lugar conductas regladas, se deben satisfacer dos roles diferentes: alguien debe actuar como siguiendo una regla y otro debe considerar que ese agente está gobernado por esa regla.⁴⁶ Entonces, desde este enfoque, ¿qué significa que alguien sigue una regla? Recuérdese que si se quiere evitar el “regulismo”, no se debe asumir que el seguimiento de planes (la práctica reglada que aquí interesa) supone siempre una apelación a la formulación de reglas explícitas. Puede haber alguna práctica implícita que suponga la adopción de la actitud correspondiente al seguimiento de planes. Si se desea evitar el “regularismo”, hay que considerar que una conducta reglada debe ser algo más que una conducta descrita meramente como regular.

Considerar una conducta como reglada es estar dispuesto a (o preparado para) imponer sanciones cuando esa conducta se desvía de una regla, sea una regla explícita o implícita. ¿Qué implica para una conducta estar reglada? La conducta que está reglada es una conducta que es considerada como reglada. Considerar que la conducta de alguien está reglada es en sí mismo un acto de reglar. A su vez, este sujeto sólo estará ligado a reglas en tanto que haya otros preparados para tratarlo como tal. No hay circularidad en la definición o comienzo de un regreso vicioso, sino un hecho sobre la normatividad: la normatividad no es reducible a la no normatividad. El espacio de las razones es creado por nuestra propensión a castigarnos y elogiarnos mutuamente de acuerdo con normas o reglas. Lo que vale como un castigo, un elogio, una regla o una norma no está constituido por algo externo al espacio de las razones. El castigo, el elogio, la regla y la norma son lo que son porque los tratamos como tales.⁴⁷ O para decirlo de otro modo, nuestras prácticas gobernadas por reglas son en todo sentido prácticas normativas. Tratar nuestras prácticas como gobernadas por reglas es tratar a sus jugadores o participantes como uno más de nosotros, y a su vez es estar preparados (o prepararse) para evaluar esa práctica críticamente, esto es, para sancionar las conductas cuando las hallamos incorrectas o desviadas.

Sin embargo, no debe pensarse que esta tesis interpretativista significa que cualquier cosa puede estar reglada en tanto que es considerada como reglada –por ejemplo, los planetas. Pueden mencionarse dos criterios para considerar una actividad como reglada. Por un lado, consideramos que alguien juega un juego, esto es, está ligado

⁴⁵ Este giro está implicado en lo que Brandom llama “fenomenalismo acerca de las normas”: “Desde una mirada ampliamente fenomenalista sobre las normas, como la que aquí será defendida, las normas están, en un sentido relevante, en el ojo del espectador; de modo que uno no puede hacer frente a la cuestión de qué normas implícitas están en juego independientemente de la cuestión de su reconocimiento en la práctica. La dirección explicativa consiste en ofrecer, en primer lugar, una explicación de la actitud práctica de tomar algo como siendo correcto según una práctica, y, en segundo lugar, explicar la condición de ser correcta de acuerdo con una práctica apelando a esas actitudes prácticas.” (Brandom, 1994, p. 25).

⁴⁶ Por supuesto, estos roles no denotan necesariamente dos personas diferentes, puesto que un mismo individuo puede ocupar alternativamente (i.e., en momentos diferentes) los dos roles. Éstos identifican las diferentes actividades y habilidades involucradas en las prácticas normativas.

⁴⁷ Véase la siguiente cita de Brandom (1994, p. 649): “El intento teórico de localizar la “fuente” de la dimensión normativa en el discurso, nos conduce de regreso a nuestras propias prácticas implícitamente normativas. La estructura de estas prácticas puede ser dilucidada, pero siempre desde el interior del espacio normativo, a saber, desde el interior de nuestra práctica normativa de dar y pedir razones. El objetivo no es reducir, sino expresar: hacer explícita la característica de la estructura en la práctica discursiva como tal.” Para una visión distinta, aunque coincidente respecto de la irreductibilidad de lo normativo a lo no-normativo, véase McDowell (1994).

a sus reglas, solamente cuando su conducta se corresponde generalmente con esas reglas. Su conducta debe aproximarse, en cierto grado, a las reglas del juego. Esta es una razón que explica por qué es poco apropiado considerar que los jugadores de truco están jugando al póquer. En segundo lugar, las sanciones son apropiadas solamente cuando tienen sentido o producen una diferencia; esto es, en sistemas que pueden ser corregidos, que aprenden de las críticas, etc. No tiene sentido considerar que los planetas siguen reglas porque en nuestra interacción con ellos no hay lugar para sanciones. Ahora bien, qué grado de desviación de las reglas está permitido antes de descalificar a una persona y dejar de considerarla un jugador o participante, es algo que variará de práctica en práctica. A continuación presentaré una forma posible en que esta discusión puede emplearse para elucidar el seguimiento de planes en el contexto de las prácticas de actuación en el mundo.

La pregunta inicial planteaba el contraste entre seguir un plan y actuar azarosamente de acuerdo con un plan. Según el enfoque interpretativo adoptado, este contraste puede tratarse a partir de esta cuestión: qué subyace al seguimiento de planes. Desde mi punto de vista, la clave para enfrentar esta cuestión puede extraerse de las siguientes palabras de Myles Brand (1984, p. 240-241): “La persona debe instanciar una estructura jerárquica de tipos de acción, y debe devenir consciente, en cierto sentido, de lo que está haciendo. La persona debe tener una representación mental de esta estructura abstracta y debe conocer que estas acciones se conforman a esta representación”. Esta apreciación sugiere que los planes guían la conducta de modo análogo a como los mapas guían nuestro desplazamiento por el espacio. Como los mapas, los planes no guían al agente mediante un acto. Por consiguiente, el guiar no es aquí una actividad que el plan ejerza sobre el agente. Por el contrario, el agente desarrolla una actividad intencional: sigue las reglas o instrucciones que conforman el plan. Para ello, estas reglas o instrucciones deben ser objeto de creencia o atención del agente. Esto significa que actuar guiado por instrucciones o reglas supone entender esas instrucciones o reglas, algo que, a su vez, implica tener creencias verdaderas sobre ellas. Por lo tanto, la etiología de una conducta, resultado de la instrucción y el seguimiento de reglas, entraña factores externos al agente: las reglas e instrucciones que conforman el plan así como su interpretación correcta. La existencia de estos factores externos sitúa al agente dentro de un espacio normativo. Desde este punto de vista, las acciones del agente pueden ser evaluadas puesto que requieren una actividad intencional del agente dirigida a aprehender su carácter normativo y a ejecutar los cursos de acción correspondientes.

El seguimiento de planes que involucran actuaciones en el mundo supone, entonces, la apelación a reglas explícitas.⁴⁸ Sin embargo, esto no sitúa a las prácticas de actuación en el mundo como blanco de críticas similares a las dirigidas a la posición etiquetada como “regulismo”. Del hecho de que las reglas explícitas no sean condiciones necesarias de todas las prácticas no se sigue que no puedan ser condiciones de algunas prácticas regladas. Además, en el caso del seguimiento de planes se satisfacen los dos criterios de la tesis interpretativista: como están en juego reglas explícitas se trata de interpretar si las conductas del agente se corresponden generalmente con esas reglas así como aplicar las correcciones o sanciones necesarias en caso de que eso no ocurra (las

⁴⁸ Imagínese la operación de cualquier artefacto de una línea de montaje para satisfacer cierto objetivo. El hecho de que las acciones del operario puedan ser corregidas supone no sólo que dicho agente sigue reglas que prescriben acciones, sino que puede seguirlas erróneamente. De allí que se trate de un hecho normativo: el operario está sujeto a la corrección.

sanciones producen realmente una diferencia en el caso de estas prácticas: seguimiento de reglas concretas de actuación).

6. Las intenciones y el razonamiento práctico: las intenciones como insumo y como resultado

En los párrafos que siguen deseo retratar de manera simple el proceso de formación de las intenciones en el contexto del razonamiento práctico. Con esto persigo un triple propósito. Por una parte, ofrecer una imagen más acabada de cómo la tradición filosófica que aquí interesa, tradición que hunde sus raíces en Aristóteles (1970b), conjuga la idea de intención con la de elección para actuar en un determinado tiempo y espacio. Por otra parte, complementar la sugerencia de que una intención es un insumo para razonamientos prácticos posteriores con la idea de que también es un producto de razonamientos prácticos precedentes. Finalmente, resaltar, mediante una revisión rápida del proceso de formación de intenciones, la idea de que ellas entrañan la generación de planes en muchos casos contingentes y parciales. Esto último servirá para enfatizar nuevamente la relación entre las intenciones y el fenómeno de los planes y la planificación, puesto que, para decirlo con palabras de Brand (1984, p. 32), “El sentido común o la psicología de sentido común parecen recomendar una noción de actividad planeada en la cual la deliberación consciente es necesaria”.

Más arriba, en la sección 2.3, afirmé que las intenciones son un insumo para razonamientos prácticos posteriores y mencioné el papel que, una vez formadas, desempeñan como aseguradoras de la coordinación personal e intrapersonal gracias a su estabilidad. Se arguyó que, por ejemplo, tener la intención de realizar cierta acción (A) guía el razonamiento sobre cómo realizar esa acción (A), a saber, el razonamiento que involucra el acoplamiento medios-fines. La intención funciona como un filtro que descarta las alternativas que el agente cree incompatible con la realización intencional de esa acción (A) y conduce al agente a empeñarse en la realización de dicha acción (A), produciendo los ajustes necesarios según las señales de éxito y fracaso recibidas durante dicho empeño. Sin embargo, las intenciones son también el resultado (output) de procesos de razonamiento práctico. Para decirlo en pocas palabras, si Valentina está deliberando sobre ir o no a la piscina esta tarde y al final decide hacerlo, la intención de ir a la piscina es lo que naturalmente llamaríamos ‘un producto típico de la deliberación de esa persona’ (Valentina, en este caso). Por consiguiente, la intuición que está en el punto de partida de esta sección puede resumirse en la siguiente aserción: las intenciones con respecto al futuro son a menudo los outputs de la deliberación.⁴⁹ Afirmación que funciona como contraparte de la anterior: las intenciones son insumos (inputs) para razonamientos prácticos. Veamos con más detalle este asunto.

⁴⁹ La afirmación menciona una restricción: las intenciones sobre el futuro son ‘a menudo’ los *outputs* de la deliberación. Si lo fuesen siempre, esto es, si decidir hacer cierta acción (A) fuese una condición necesaria para intentar esa acción (A), entonces habría que esforzarse por desarrollar una explicación del intentar sobre la base de una explicación previa del decidir. La restricción se debe a que no todas las intenciones se forman como resultado de una decisión (por ejemplo, en muchos contextos respondemos preguntas, atendemos el llamado a nuestra puerta o respondemos el teléfono sin embarcarnos conscientemente en un proceso de toma de decisión). La deliberación consciente no es el único mecanismo de formación (o adquisición) de intenciones: véase, por ejemplo, Mele (1992, p. 141) y Brand (1984, especialmente pp. 129-130). Sin embargo, no discutiré aquí esos otros mecanismos.

¿Qué sucede cuando un agente delibera y elige? Según el esquema clásico aristotélico de la deliberación práctica, en su punto de partida se encuentra la representación de un fin u objetivo deseado. El segundo paso está dado por la creencia del agente de que está en su poder realizarlo, o dicho de otro modo, llega a creer, como resultado de la deliberación, que una cierta clase de acción, cuya realización está en su poder, serviría para producir el objetivo deseado (Aristóteles, 1970, 1112b, p. 37). La cuestión que plantea la deliberación práctica es, en cierto sentido, la siguiente: ¿qué clase de acción (o plan de acciones) a mi alcance (i.e. en mi poder) producirá (o conformará el primer paso para producir) el objetivo buscado? Esta pregunta se resuelve cuando efectivamente se elige realizar una acción de la clase que produce el objetivo. De este modo, la deliberación contribuye a la formación de una intención sobre el futuro (i.e. nos representamos estados de cosas no actuales hacia los cuales nuestra actitud entraña un compromiso) en el contexto de una decisión o elección.⁵⁰

En esta interpretación del razonamiento práctico, lo importante, para decirlo directamente, es que la conclusión de la deliberación no es la única cosa que resulta razonable hacer prácticamente. Tampoco es la única cosa que siempre será razonable hacer (Elster, 1997). No advertir estas dos cuestiones, como bien señaló Anscombe (1963, p. 61), conduce a interpretaciones “insanas” de la propuesta aristotélica. Una interpretación “insana” supone que si el razonamiento práctico ha de explicar la acción, sus premisas deben ser universales y su conclusión individual debe ser un caso que caiga bajo la conclusión universal que se sigue deductivamente de las premisas. Sin embargo, interpretar de manera saludable el razonamiento práctico no significa despojarlo de su carácter de pieza deductiva; por el contrario, sólo implica reconocer que existen distintas conclusiones prácticas, puesto que hay otras maneras igualmente satisfactorias de realizar un objetivo. Por consiguiente, la elaboración del plan de acción (i.e. el proceso de formación de la intención sobre el futuro), cuya realización deviene con la elección, es un curso de acción entre otros (cursos satisfactorios de acción) posibles.⁵¹ El agente elige un curso de acción (o traza un plan) sobre la base de examinar las características concretas, contingentes y particulares de su situación, o dicho de otro modo, la elección que realiza el agente de los productos de su deliberación tiene que ver con los cursos de acción que es capaz de representarse y seguir para lograr el objetivo. Entonces, una interpretación saludable (i.e. no insana) del razonamiento práctico dice que no hay un plan óptimo fuera

⁵⁰ Adviértase que una acción no es una actitud proposicional, puesto que no es una pintura del mundo, sino una contribución causal al mundo. Sin embargo, formar una intención para realizar una acción implica tener una actitud proposicional. Por consiguiente, el objetivo del razonamiento práctico es alcanzar una conclusión cuya aceptación constituya la formación de una intención para actuar. Brand (1984) retrata muy bien esto último en relación con Aristóteles cuando dice: “El espíritu de la afirmación de Aristóteles, pienso, puede ser capturado considerando [...] la intención como siendo un evento representado por la conclusión de un [práctico] silogismo. La intención causa la acción, teniendo en cuenta que el ambiente es amistoso [...] Si la acción prospectiva es compleja, la intención conlleva un plan para lograrla” (p. 129).

⁵¹ Es posible argumentar que se puede desarrollar una pieza de razonamiento práctico que lógicamente requiere una elección y, no obstante, no realizar esa elección. En este punto conviene señalar que la actitud hacia el objetivo que se pretende realizar no es un mero deseo. Por el contrario, se trata de un querer para el cual se pretende una realización. Dicho de otra manera: puedo intentar realizar un objetivo sin tener mucha idea de cómo realizarlo, pero no puedo seguir intentándolo a menos que continúe deliberando hasta alcanzar una conclusión sobre cómo realizarlo. Por otra parte, recuérdese la sugerencia de Austin (1975) de que las intenciones no pueden ser demoradas.

o independientemente de las circunstancias del agente.⁵² Los planes deben adecuarse a sus capacidades y a las circunstancias de la deliberación. Desde este punto de vista, los planes se vuelven en muchos casos esquemas contingentes de acción.⁵³ El curso de la deliberación entraña un proceso de planificación que está sometido a cuestiones prácticas contingente y particulares. A través de él se forma una intención cuya realización supone la ejecución de una elección.⁵⁴

7. La acción intencional en función de la noción de plan: aproximación a su estructura básica

Este estudio comenzó con una intuición recogida por nuestros usos lingüísticos ordinarios: actuar intencionalmente entraña de manera general actuar según una intención. Si un individuo actúa intencionalmente hay algo que ese individuo intenta hacer; por tanto, existe un vínculo relevante entre intentar hacer algo y hacer ese algo intencionalmente. Desde este punto de vista, aquello que hace que la realización de una acción (A) por parte de un agente (S) sea una realización intencional, es la relación de esa acción (A) con una intención relevante: la intención de hacer (A). Dentro del marco de esta intuición, el análisis del contenido de la noción de intención se transformó en una cuestión prioritaria, porque es precisamente ese contenido y las relaciones que mantiene con cierta acción lo que vuelve a ésta intencional. En términos generales, las intenciones cumplen una doble función determinativa: cualifican la mente del agente y sus acciones. Por consiguiente, el análisis de la naturaleza y contenido de la intención no puede emanciparse de una propuesta teórica que dé cuenta de cómo se relacionan las intenciones qua estados mentales con las acciones realizadas intencionalmente o con cierta intención.

Para realizar esta tarea, se construyó una propuesta filosófica entroncada con las elaboraciones aristotélicas sobre el razonamiento práctico. Según esta propuesta, el enlace entre la intención qua estado mental y la acción realizada intencionalmente se comprende haciendo hincapié en la noción de intención sobre el futuro. El análisis de los contenidos de este caso central de intención puso en evidencia la dimensión planificadora de la intención: la intención de realizar cierta acción entraña la planificación sobre cómo

⁵² Quizá, no debería hablar de plan óptimo sino de plan satisfactorio; no obstante, la idea de un plan óptimo según las propias circunstancias del agente se acerca mucho a la idea de plan satisfactorio. Esta última sugerencia, por otra parte, no implica restarle racionalidad a la elección del agente; como bien señala Elster (1997), “racionalidad y optimalidad no son sinónimos. La racionalidad implica en sentido mínimo la consistencia de metas y creencias” (p. 66)

⁵³ Que el agente se decida por un curso de acción entre otros no implica que no reconozca los restantes cursos de acción posibles. Los aristotélicos medievales empleaban la distinción entre elección y consentimiento para describir esta situación (véase Tomás de Aquino, 1988, I y II, 15, 3, p. 162-3). Para este autor, hay muchas conclusiones prácticas sobre las que el agente es necesariamente consciente pero que no necesariamente elige. En sus palabras, “Pues puede suceder que mediante el consejo se encuentren varias cosas que conducen al fin y, si todas ellas agradan, se consiente con todas ellas; pero de esta cantidad de cosas que agradan escogemos sólo un eligiendo” (1988, I y II, 15, 3, p. 162)

⁵⁴ Recordemos que si bien un plan de acción puede formarse en el curso de un razonamiento práctico sobre cómo alcanzar cierto objetivo, no necesita formarse siempre de esa manera. Aquí se trata de decir que no todos los planes de acción son resultado de un razonamiento práctico completo. Hay muchos planes fragmentarios, asociados a circunstancias concretas, cuya formación no llega a ser completamente consciente. Sobre este último punto, véase, entre otros, Brand (1984). Por otra parte, no siempre que ejecuto un plan ejecuto una elección: no elijo cada mañana el plan para ir caminando a la biblioteca; de forma prácticamente inconsciente sigo una rutina almacenada en mi memoria.

producir esa acción, el compromiso efectivo con sus pasos preliminares y, finalmente, la conducción (o guía) adecuada de su realización. Esta elaboración no sólo permitió anclar la intuición filosófica de que la intención implica en un sentido fuerte la planificación, sino también poner las bases para transformar el análisis de la intención en el estudio de la formación y estructura de los planes de acción. Los resultados de este estudio permiten aproximarnos ahora a la caracterización de la estructura básica de la acción intencional acentuando la noción de plan de acción.

La estructura básica de la acción intencional puede caracterizarse de la siguiente manera: la acción intencional es la acción actuada (o realizada) siguiendo un plan. De este modo, un agente (S) realiza intencionalmente una acción (A) durante cierto tiempo (t) sólo si dicho agente (S) tiene un plan (P) durante ese tiempo (t), tal que su acción (A) está incluida en el plan (P), y sigue ese plan (P) cuando realiza esa acción (A).

Esta caracterización, por un lado, subraya el hecho general de que actuar según una intención implica contar con una representación de la clase de acción que se ha de realizar y, por otro, señala que se actúa intencionalmente sólo cuando se sigue un plan y la acción efectuada forma parte efectiva de ese plan. Sin embargo, si se toman en cuenta los desarrollos de las secciones anteriores, esta caracterización requiere ser precisada en un número significativo de aspectos.

En primer lugar, los planes, implicados en las intenciones, son en general planes parciales.⁵⁵ Esto es, los planes *qua* estados mentales no necesitan involucrar representaciones completas de los cursos posibles de acción futura; por el contrario, en los casos de planes *qua* estados mentales, éstos son estructuras relativamente abiertas e incompletas. En virtud de estas propiedades, los planes pueden ser moldeados por los agentes según las circunstancias, para ajustarse a ellas en tiempo real. Si los planes no se modificaran, ajustaran y combinaran entre sí, los agentes no podrían alcanzar objetivos complejos ni coordinarse individual y colectivamente a partir de recursos limitados.⁵⁶ Por consiguiente, gracias a que los planes son parciales, los agentes humanos pueden razonar a partir de planes previos sobre planes futuros, a partir de planes generales sobre planes específicos o a partir de planes ya trazados sobre los respectivos subplanes en su condición de medios y pasos preliminares. Aún más, en muchas situaciones cotidianas el agente elabora un plan general que se va rellenando con subplanes a través de un procedimiento por defecto (default), esto es, en función de la interacción entre las opciones propias del mundo y los cursos de acción almacenados en el agente como rutinas. Por ejemplo, imagínese la siguiente situación: Valentina, una joven cinéfila, tiene la intención de ir al cine esta noche a ver la última película de los hermanos Cohen. Sin embargo, no planifica cada uno de los cursos de acción que ha de emprender cuando llegue al cine. Una vez allí, se activan por defecto subplanes almacenados en su memoria concernientes a cómo obtener su entrada, cuál es el procedimiento de ingreso a la sala,

⁵⁵ Recuérdese la metáfora de Austin de la “lámpara de minero”, una lámpara que “no ilumina todos mis alrededores” (Austin, 1975, p. 259).

⁵⁶ Recuérdese que esta dinámica de modificación, ajuste y combinación hay que apreciarla sobre el trasfondo de la inercia que presentan los planes. Al buscar la coordinación y acción efectiva los agentes trazan planes parciales que, si no surge información nueva relevante o alguna contingencia propia del mundo como problema, tienden a retener. En cualquier caso, la tensión entre la estabilidad de un plan parcial y su revocabilidad implica el análisis de la sensibilidad del agente para con los problemas y oportunidades que ofrece el medio ambiente, puesto que la realización de un plan supone información incompleta sobre el futuro.

etc. –es decir, se activan subplanes que de algún modo le ayudan a estructurar su medio ambiente y controlar racionalmente su conducta.⁵⁷

En segundo lugar, así como no todo lo que figura en los planes es algo que se intenta, existen acciones que pueden ser intencionales sin formar parte explícitamente del plan previamente representado. La razón es muy sencilla: el boicot de las circunstancias del mundo y nuestras limitadas capacidades hacen que resulte prácticamente imposible disponer anticipadamente de los medios adecuados para enfrentar todas las posibles contingencias futuras. Esto posibilita realizar una consideración sobre un aspecto importante de la caracterización de la estructura de la acción intencional: se actúa intencionalmente sólo cuando las acciones realizadas forman parte de un plan (el plan que efectivamente se sigue). Pero, entonces, ¿qué significa decir que hay acciones que pueden ser intencionales sin formar parte del plan concebido explícitamente? En términos generales, significa que, aunque las acciones son intencionales por formar parte de un plan determinado, cada acción intencional no necesita ser planificada –i.e. no necesita ser explícitamente concebida. En términos particulares, significa que una acción puede heredar la propiedad de ‘ser intencional’ de un patrón general de actividad (Brand, 1984; Bratman, 1987).

Imagínese la siguiente situación: Fortunato, fuertemente influido por los escritos de Bertold Brecht, forma la intención de atracar una determinada sucursal del Banco Francés en Buenos Aires; planifica ingresar algunos minutos después de su apertura en la sucursal de avenida Coronel Díaz, entre las 10:35 y las 11:00, dirigirse a la oficina del gerente de la sucursal, amenazarlo con un revólver y ordenarle que instruya al empleado restante para que deposite el dinero de la caja fuerte en un maletín; luego amarrará a ambos empleados en la oficina del gerente, les echará gas paralizante y se dará a la fuga. Sin embargo, contra el pronóstico basado en pasadas y reiteradas observaciones del estado de la sucursal bancaria durante esa franja horaria, Fortunato se encuentra con que hay un cliente en el banco. Finalmente, roba el banco pero no de la manera planificada, sino reduciendo tanto al cliente como al empleado del mostrador para ordenar al gerente, amenazándolo con un revólver, el depósito del dinero en su maletín y dándose a la fuga. Estas acciones no formaban parte del plan; por el contrario, son acciones que Fortunato agregó de manera imprevista para satisfacer con éxito su objetivo. No obstante, esas acciones son intencionales, y lo son porque, aunque no estén explícitamente planificadas, se inscriben en (o pertenecen a) un patrón de actividad intencional: el plan general del robo de esa sucursal bancaria. Para decirlo con palabras de Brand (1984, p. 28): “La intencionalidad concurre a las acciones focales en un plan. Estas acciones focales, consideradas juntas, forman un patrón de actividad; y otras acciones son intencionales en virtud de ser parte de este patrón de actividad”. En consecuencia, una acción puede ser sustituida por otra cuando la acción planeada no puede llevarse a cabo, o también puede improvisarse cuando forma parte integral de una actividad planificada. Y en ambos casos, se trataría de una acción intencional.⁵⁸

⁵⁷ Schank y Abelson (1977, p. 41 y 72) distinguen entre planes y guiones. Los guiones, a diferencia de los planes, entrañarían secuencias de acciones estereotipadas y predeterminadas que concretan una situación bien definida. Por consiguiente, la idea de subplanes que se activan por defecto en situaciones conocidas, tendría que ver con estos guiones. Por otra parte, la razón de que se los denomine ‘guiones’ reside en que se organizan según una estructura narrativa. Agradezco a Fernando Broncano por haberme señalado oportunamente la importancia de este punto.

⁵⁸ Brand (1984, especialmente p. 28-31) precisa esta observación codificando los tipos específicos de acción que, según nuestro sentido común, heredarían su intencionalidad del patrón general de actividad. Estos

En tercer lugar, sobre la base de esta caracterización se puede retratar lo que sería la formulación mínima de la estructura básica de la acción intencional. Un caso de formulación mínima (o límite) es un caso de acción intencional en el que el plan involucra una representación rudimentaria o elemental de la acción prospectiva. O dicho en los términos de Austin (1975, p. 249-271) la idea, concepción o noción que tiene el agente de lo que lleva entre manos (o está haciendo) es de una extensión mínima: la idea de hacer algo por el mero hecho de hacerlo y no como medio para producir otra cosa (i.e., hacerlo por sí mismo).⁵⁹ El contenido de la intención reflejaría sólo la meta u objetivo de la acción, puesto que fin y medios coincidirían. A partir de esta fórmula básica, se construyen los casos más complejos. Estos casos involucran acciones intencionales mediadas. La composición de las acciones intencionales mediadas supone la articulación instrumental del contenido de la intención. Es decir, el contenido de una intención de hacer (A) se describe instrumentalmente como, por ejemplo, la realización de cierta acción (B) por medio de cierta acción (C). Producir (C) es un acto instrumental del agente para producir (B). En estos casos, el plan del agente necesita de los medios representacionales adecuados para articular la complejidad de las relaciones instrumentales en juego.⁶⁰

En cuarto lugar, adviértase que si bien la caracterización dada no recoge, dentro de la estructura básica de la acción intencional, ninguna relación entre el plan, sus acciones y las creencias propias del agente, eso no significa que no haya ninguna. En la sección 2.3 se menciona en una nota a pie de página que un agente puede intentar hacer (A) y realizar intencionalmente (A) sin tener la creencia de que realizará exitosamente esa acción (A). Sin embargo, no hay que malinterpretar la fuerza conclusiva de esta afirmación. Que el agente no tenga la creencia de que realizará con éxito la acción respectiva, no significa que no tenga en absoluto la creencia de que realizará esa acción con independencia de su éxito práctico.⁶¹ En consecuencia, la afirmación anterior no debe leerse como una aseveración a favor de la ausencia de lazos entre las creencias del agente y sus planes, intenciones y acciones. No obstante, ¿cómo podría intervenir la creencia en la acción intencional? ¿Cuál sería su papel?

tipos se encontrarían especificados por tres principios relativos a las acciones: el principio de sustitución, improvisación e interrupción. En cualquier caso, lo que importa es que una acción retiene su condición de acción intencional siempre y cuando sea parte de un plan, incluso si ese plan no se realiza de la manera originalmente concebida.

⁵⁹ Por ejemplo, el plan involucrado en la intención de levantar la pierna. Por su parte, Goldman (1970) llama a esta clase de acción intencional: “acción intencional básica” (p. 24).

⁶⁰ Adviértase que en estos casos el plan hace explícita referencia al contenido instrumental de la intención. Mosterín (1978, p. 151 y 159) emplea los términos “acción mediada” y “acción compuesta” para diferenciar acciones según el contenido de la intención (en su terminología, según la idea que se quiere realizar “ φ ”). Así, una acción mediada es una acción donde el contenido de “ φ ” involucra una relación causal, por ejemplo, “ π causa φ ”; mientras que una acción compuesta es una acción donde el contenido de “ φ ” involucra una cadena causal (una extensión del componente causal de la intención), por ejemplo, “ π_1 causa π_2 causa ... π_n ”.

⁶¹ En este punto podría distinguirse entre creencia y confianza (o expectativas); es decir, uno podría tener la creencia de que realizará (A), pero no la confianza (o expectativa) de que hará (A) exitosamente. Quizá, podríamos formular negativamente la inclusión de la creencia en la acción intencional: un agente (S) intenta (A) sólo si (S) no cree en el momento apropiado que no realizará (A). Véase, entre otros, Audi (1993), Mele (1989 y 1992).

La creencia es una actitud proposicional diferente de la intención; satisface otros roles funcionales y, por tanto, no participa de su contenido.⁶² Además, la creencia es una entidad distinta de la acción (creer que *p* no es análogo a hacer *p*). En tanto que hecho empírico, la creencia precede y/o es contemporánea de la acción. La creencia admitiría un test contrafáctico como el siguiente: si el agente no hubiese creído lo que efectivamente creía (por ejemplo, que cierta acción era un medio necesario para la realización del objetivo intentado), no habría (o podría no haber) actuado de la forma en que realmente lo hizo. De esto se sigue que la creencia determina en algún sentido la acción. Por consiguiente, la creencia puede ser, al menos, una parte no redundante de una condición necesaria para la realización intencional de la acción en cuestión. De hecho, cuando un agente (*S*) intenta realizar una acción (*A*) está dispuesto a realizar no aquellas acciones que son de hecho necesarias para la realización de (*A*), sino aquellas que (*S*) cree que son necesarias para realizar intencionalmente (*A*). Por ende, el agente intencional posee creencias que tratan sobre asuntos fácticos, relaciones causales entre medios y fines, etc. Desde esta perspectiva, las creencias contribuirían al basamento y coordinación de la acción, suministrando planes basados en el descubrimiento de relaciones de mediación causal entre acciones.⁶³

En la sección siguiente discutiré brevemente el caso de la acción intencional espontánea. Este caso es especialmente importante puesto que representa un obstáculo para la consideración de la intención sobre el futuro como caso central de intención. Sin embargo, mi propósito no será realizar un examen completo de la noción de intención en la acción; por el contrario, sólo intentaré poner de manifiesto que este caso no es únicamente un caso problemático.

8. Un problema para la intención sobre el futuro: la intención en la acción o el caso de la acción intencional espontánea

Dicho en pocas palabras, la objeción de Searle (1983) ataca la intuición filosófica de que es posible transformar el estudio de la intención en un estudio sobre la planificación.⁶⁴ Según este autor, no se justifica un proceder de esta clase. Citaré en extenso a Searle con el propósito de retratar su objeción:

“[M]uchas de las acciones que uno realiza, las realiza espontáneamente, esto es, sin formar, consciente o inconsciente, ninguna intención previa a la realización de esas acciones. Por ejemplo, supongamos que estoy sentado en

⁶² Por ejemplo, producir y guiar adecuadamente la conducta es algo que realiza la intención; sin embargo, no depende de que el agente esté en cierto estado doxástico, o dicho de otro modo, es independiente de la verdad de las creencias del agente. Véase Brand (1984, p. 152).

⁶³ Por ejemplo, intento hacer (*A*), pero descubro que hacer (*B*) resultaría en la producción de (*A*). Por ende, planifico hacer (*B*) como medio para hacer (*A*) –véase, entre otros, Mosterín (1978, p. 151). Por otra parte, nótese que los planes deben ser consistentes con las creencias del agente, esto es, debe ser posible ejecutar los planes con éxito en un mundo en el cual las creencias del agente sean verdaderas.

⁶⁴ La etiqueta ‘objeción de Searle’ merece una aclaración. Por una parte, si bien se presenta esta objeción en referencia a la persona de Searle, éste no ha sido el único en proponer la idea de intención en la acción o de acción intencional espontánea; Brandom (1994), por ejemplo, también lo ha hecho. Sin embargo, Searle (1983) es quien ataca explícitamente la idea de que no puede tratarse la intención sobre el futuro como caso central de intención, puesto que hay casos de acción intencional que no entrañan intención alguna sobre el futuro.

mi sillón reflexionando sobre un problema filosófico y, de repente, me paro y comienzo a dar vueltas por la habitación. Pararme y dar vueltas por la habitación son claramente acciones intencionales; sin embargo, para llevarlas a cabo no necesité formar una intención previa de llevarlas a cabo. En ningún sentido necesito tener un plan para levantarme y dar vueltas por la habitación. Como muchas cosas que hago, meramente las hago; solamente actúo” (1983, p. 84).

Su objeción puede reconstruirse del siguiente modo: (a) existe una casuística que muestra que hay acciones intencionales espontáneas. Por tanto, (b) no todas las acciones intencionales son acciones que resultan de llevar a cabo un proceso deliberativo, esto es, no todas estas acciones son acciones planeadas.⁶⁵ Estos dos pasos, (a) y (b), sugieren (c): hay intenciones que no involucran planes. A estas intenciones, Searle (1983, pp. 79-111) las denomina “intenciones en la acción”. El corolario de esta clase de objeción podría sintetizarse en una afirmación de este estilo: en ciertas acciones hay intenciones en la acción (o dicho negativamente: no hay intenciones previas) y éstas no suponen en ningún sentido un plan.⁶⁶ De esto se sigue que la intención sobre el futuro no podría considerarse un caso central de intención.

A esta objeción se puede responder de diferentes maneras. En primer lugar, se puede responder socavando la afirmación misma de Searle. Un procedimiento de esta clase podría articularse según un doble movimiento argumentativo, conformando un ataque sobre dos frentes filosóficos distintos. Por una parte, argumentando que Searle considera erróneamente que la “intención en la acción” es una subclase del tratar (“trying”).⁶⁷ Las intenciones no son una subclase del tratar. No lo son en realidad por una razón sencilla a primera vista: una acción es una contribución al mundo, no una figura del mundo. Si las intenciones fuesen una subclase del tratar, podrían ser consideradas en cierto sentido acciones y, en consecuencia, contribuciones al mundo.⁶⁸ No obstante, una intención no es una contribución al mundo. Por consiguiente, las intenciones no son

⁶⁵ Adviértase que la idea de Searle es que hay acciones intencionales espontáneas. Por tanto, su objeción no es una objeción hacia la restricción de la acción a la acción intencional; es decir, Searle no está defendiendo la intuición de que hay acciones imprevistas (no planeadas), espontáneas o impulsivas (esto es, acciones que no están precedidas por un razonamiento práctico). Esto no necesita ser defendido, puesto que de hecho las hay; por el contrario, la idea de Searle es que hay acciones intencionales espontáneas; y esto es lo que aquí está en cuestión.

⁶⁶ Quizá, una interpretación caritativa de la intención en la acción sea considerar que se asemeja a los mecanismos de guía y control de la acción implicados en las intenciones. Esto supondría reconstruirla según estos mecanismos, pero difícilmente Searle estaría de acuerdo. Para una sugerencia de este estilo, véase Brand (1984, especialmente capítulo 7).

⁶⁷ La sugerencia de que la idea searleana de intenciones en la acción es una subclase del tratar está desarrollada en Mele (1994, capítulo 10, especialmente p. 183 y ss.): “una intención en la acción es efectiva –resulta en movimientos parcialmente constitutivos de una acción intencional pertinente– cuando un agente tiene éxito en hacer lo que está tratando de hacer, *porque* está tratando de hacerlo”. El argumento de Mele contra Searle es que esta idea no descarta la presencia efectiva de las intenciones sobre el futuro. La ansiedad de Mele reside en que propuestas como las de Searle puede dar lugar a la idea de que “alguien puede intencionalmente hacer A sin intentar hacer algo en absoluto” (1994, p. 194). Para una crítica más amplia de la posición de Searle, véase también Mele (1989, pp. 101-106).

⁶⁸ La premisa subyacente es que el tratar, a diferencia de la intención, es una acción; esto puede advertirse rápidamente cuando se piensa que una intención de hacer (A) puede en parte ser una disposición a tratar de hacer (A). La intención figuraría en la etiología causal del tratar, mientras que este último compondría la acción misma.

acciones; ni siquiera son acciones parciales. De esto se sigue que, si estas intenciones son una subclase del tratar, entonces no hay intenciones en la acción.

Por otra parte, argumentando que no hay “intención en la acción” en la acción intencional espontánea, puesto que no hay una distinción de clase entre una intención de producir en el futuro cierto objetivo a través de un plan y una intención de hacer algo aquí y ahora. Este movimiento argumentativo se aseguraría señalando que el plan que guía una acción no necesita ser un plan completo y detallado; aún más, como quedaba reflejado en el tratamiento de la idea austiniana de plan (sección 2.3), el plan puede tener un sentido mínimo: ser sólo lo que el agente piensa que está haciendo, independientemente de si ese pensamiento es o no expresado.⁶⁹

En definitiva, este doble movimiento argumentativo socavaría la objeción searleana a través de un efecto de pinzas con el siguiente resultado: por una lado, asentando que la idea de intención en la acción no es una subclase del tratar; por otro lado, mostrando que, si se decide analizar la noción de intención en la acción según la noción de intención, se puede atribuir al mismo Searle un error conceptual, puesto que el concepto de intención está relacionado por definición con la idea de futuro. Para decir esto último de otro modo, como bien ha señalado O’Shaughnessy (1980), las intenciones podrían dirigirse hacia actos ubicados completamente en el presente sólo a condición de que esos actos carecieran de extensión. En consecuencia, las intenciones dirigidas hacia actos que están realizándose se encuentran siempre orientadas hacia el sector de esos actos que permanece en el futuro. La conclusión es que la noción misma de intención no se ajusta al modo en que Searle entiende las intenciones en la acción. Por consiguiente, resultaría difícil emplear justificadamente la noción de intención en la acción.⁷⁰

⁶⁹ Para decirlo en términos de Austin (1975), un plan puede ser sólo esa idea que tiene el agente de lo que lleva entre manos; una idea de la que el agente es consciente. Por otra parte, también se podría argumentar a favor de la ausencia de la correspondiente distinción de clase entre estas intenciones, señalando con Bratman (1987) que en el caso de estas acciones espontáneas searleanas, en realidad el agente se encuentra ejecutando “some long-standing personal policy” (p. 126). De allí que “the intention involved in spontaneous intentional action need not itself be present-directed, but may rather be a long-standing general intention, one that applies to the present case” (p. 126).

⁷⁰ Creo que parte de la ansiedad que ha conducido a ciertos filósofos a defender a rajatabla la idea de intención en la acción tiene que ver con el análisis del caso de la acción moral. La relación de la cuestión moral con el problema de la intención es un asunto que no sólo viene de lejos, sino que además está plenamente justificado, puesto que la pregunta “¿Cuáles son los criterios para interpretar un acto como intencional?” puede ser sustituida en ciertos contextos de investigación filosófica por esta otra “¿Qué debe hacer una persona para ser considerada propiamente un agente?” Velleman (1989), por ejemplo, presenta un tratamiento de la intención y la acción intencional en el marco de esta última interrogación. O dicho de otro modo, los compromisos morales son una especie de compromisos prácticos y actuar por los compromisos morales mismos –i.e. en nuestra condición propia de agentes– es ciertamente un caso de acción intencional. La situación generadora de ansiedad es la siguiente. Imaginemos que nos preguntamos: “¿Cuál es el papel (rol) motivacional de un compromiso moral en el caso de la acción moral?”. Pues bien, diferentes situaciones merecen diferentes descripciones. En algunos casos, la deliberación moral y la acción pueden semejar la aplicación de una regla explícita. Una persona se enfrenta a una situación que evoca una regla de conducta ya formulada – un precepto o una ley moral– y juzga lo que la regla tiene que decir sobre esa situación. En otros casos, una persona puede simplemente percibir que la situación evoca una cierta clase de acción. Como ha enfatizado McDowell (1978, 1994), no debemos ver todo el razonamiento moral bajo el modelo regla-caso. Entonces, hay al menos dos tipos de acción moral. Una persona virtuosa que actúa por la regla misma podría actuar con una intención previa, decidir que la regla sugiere la acción siguiente y entonces actuar sobre la base de esa decisión, o podría actuar con una intención en la acción tal que su acto mismo expresara su juicio de que la regla se aplica en este caso. De igual modo, una persona podría percibir que una situación evoca una cierta respuesta, reconocer esa

En segundo lugar, suponiendo que los argumentos anteriores resultaran poco concluyentes, aún se podría responder a la objeción señalando que la casuística que menciona Searle (los casos de acción intencional espontánea) no es relevante en el contexto del análisis de la acción intencional; por el contrario, dentro de este ámbito, el caso central de intención es la intención sobre el futuro.

9. Agradecimientos

Este trabajo no podría haberse realizado sin las oportunas sugerencias a versiones anteriores de Jesús Vega Encabo y Fernando Broncano. Al mismo tiempo agradezco los constructivos comentarios recibidos por una evaluadora anónima. Este trabajo fue elaborado como parte de los siguientes proyectos: FFI2013-45659-R y PICT 2013-1419.

10. Referencias

- Adams, F. (1986). Intention and Intentional Action: The Simple View, *Mind and Language* 1, 281-301.
- Alston, W. (1986). An Action-Plan Interpretation of Purposive Explanations of Action. *Theory and Decision*, 20, 275-299.
- Anscombe, G. E. M. (1963). *Intention*. Ithaca: Cornell University Press.
- Anscombe, G. E. M. (1981). *Collected Papers* (3 vols.). Oxford: Blackwell.
- Anscombe, G. E. M. (1983). The Causation of Action. En C. Ginet y S. Shoemaker (eds.) (1983), *Knowledge and Mind*, Oxford, Oxford University Press, 174-190.

percepción, y entonces actuar sobre la base de ese reconocimiento, o su percepción podría simplemente conducirlo a actuar, sin la intervención de una intención previa. Esta taxonomía parcial de la acción moral solamente bosqueja la superficie de un asunto espinoso y profundo. No obstante, es suficiente para indicar cuál es la fuente de ansiedad: la creencia de que si no se defiende la idea de intención en la acción, entonces el razonamiento moral caería únicamente bajo el modelo regla-caso. Sin embargo, creo que es posible eliminar esta ansiedad. Si se advierte que la noción de intención previa o intención sobre el futuro posee también un sentido mínimo que puede ser sólo lo que el agente piensa que está haciendo, independientemente de si este pensamiento es expresado, la ansiedad filosófica podría eliminarse. Adviértase que en todos los casos recién mencionados, el agente moral está motivado por sus compromisos morales (una especie de compromiso práctico). Independientemente de si esto toma la forma de una intención previa o una intención en la acción, de si el agente apela a una regla respecto de la cual está comprometido o simplemente percibe la situación como demandando algún acto suyo, la toma de conciencia de sus compromisos morales es causalmente eficaz en la economía conductual del individuo. Es decir, en esas acciones intencionales morales, una persona responde a su toma de conciencia de sus respectivos compromisos morales. Y esto puede contraponerse al caso de una persona que actúa de acuerdo con sus compromisos morales sin ser consciente de ellos, una persona que simplemente hace circunstancialmente la cosa correcta, y respecto de la cual dudaríamos si hizo intencionalmente la cosa correcta (si su acción fue o no intencional bajo alguna descripción relevante). Por el contrario, nos encontraríamos tentados de decir que no ha actuado intencionalmente, esto es, que no ha actuado según el compromiso respectivo. Sin embargo, el contraste interesante entre actuar de acuerdo con una toma de conciencia de los respectivos compromisos morales u ocasionalmente de acuerdo con ellos (sin toma de conciencia), sobrevive dentro del contexto del sentido mínimo que puede revestir la idea de intención previa o intención sobre el futuro. Por tanto, no habría por qué recurrir a la idea de intención en la acción para recoger aquellos casos que no parecen involucrar directamente el modelo regla-caso. Si bien este asunto amerita un tratamiento más extenso y complejo, creo que la distinción que realmente importa puede albergarse sin problemas dentro de la idea de intención que nos interesa. Por consiguiente, no habría razones compulsivas para verse obligado a defender la idea de intención en la acción en el contexto del análisis del caso de la acción moral.

- Aristóteles, (1970). *Ética a Nicómaco*. Madrid: Instituto de Estudios Políticos.
- Audi, R. (1977). *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press.
- Audi, R. (1991). Intention, Cognitive Commitment, and Planning. *Synthese*, 86, 361-78.
- Audi, R. (1993). *Action, Intention, and Reason*. Ithaca-London: Cornell University Press.
- Austin, J. L. (1970). *Philosophical Papers*. London: Oxford University Press.
- Austin, J. L. (1975). *Ensayos filosóficos*. Madrid: Revista de Occidente.
- Baier, A. (1970). Act and Intent. *Journal of Philosophy*, 67, 648-58.
- Bennett, J. (1966). Real. *Mind*, LXXV, 504-507.
- Brand, M. (1982). Cognition and Intention. *Erkenntnis*, 18, 165-87.
- Brand, M. (1984). *Intending and Acting. Toward a Naturalized Action Theory* Cambridge, Mass.: MIT Press.
- Brandom, R. (1994). *Making it Explicit. Reasoning, Representing and Discursive Commitment*. Cambridge, Mass.: Harvard University Press.
- Bratman, M. (1981). Intention and Means-End Reasoning. *The Philosophical Review*, 90, 252-265.
- Bratman, M. (1983). Taking Plans Seriously. *Social Theory and Practice*, 9, 271-287.
- Bratman, M. (1984). Two Faces of Intentions. *Philosophical Review*, 93, 375-405.
- Bratman, M. (1987). *Intentions, Plans, and Practical Reason*. Cambridge, Mass.: Harvard University Press.
- Bratman, M. (1992). Shared Cooperative Activity. *Philosophical Review*, 101, 327-41.
- Bratman, M. (1993). Shared Intentions. *Ethics*, 104, 97-113
- Bratman, M. (1999). I Intend That We J. En M. Bratman, *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press, 142-261.
- Bratman, M. (2000). Reflection, Planning, and Temporally Extended Agency. *Philosophical Review*, 109, No. 1, 35-61.
- Bratman, M. (2007). *Structures of Agency*. Oxford: Oxford University Press.
- Bratman, M., D. Israel y M. Pollack (1988). Plans and Resource-Bounded Practical Reasoning. *Computational Intelligence*, 4(4), 349-355.
- Castañeda, H. N. (1975). *Thinking and Doing*. Dordrecht: Reidel.
- Davidson, D. (1963). Actions, Reasons, and Causes. *Journal of Philosophy*, 60, 685-700.
- Davidson, D. (1967). The Logical Form of Action Sentences. En N. Rescher (ed.), *The Logic of Decision and Action*. Pittsburgh, Pa.: University of Pittsburgh Press.
- Davidson, D. (1970). How is Weakness of the Will Possible? En J. Feinberg (ed.), *Moral Concepts*, Oxford: Clarendon Press.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Clarendon Press.
- Donagan, A. (1987). *Choice*. London-New York: Routledge & Kegan Paul.
- Elster, J. (1997). *El cambio tecnológico. Investigaciones sobre la racionalidad y la transformación social*. Madrid: Gedisa.
- García, J. L. A. (1990). The Intentional and the Intended. *Erkenntnis*, 33, 191-209.
- García, J. L. A. (1991). On the Irreducibility of the Will. *Synthese*, 86, 349-360.

- Goldman, A. (1970). *A Theory of Human Action*. Englewood Cliffs: Prentice-Hall.
- Harman, G. (1976). Practical Reasoning. *Review of Metaphysics*, 79, 431-463.
- Harman, G. (1986). *Change in View*. Cambridge: Mass., MIT Press.
- Holton, R. (2009). *Willing, Wanting, Waiting*. Oxford: Oxford University Press.
- Hornsby, J. (1980). *Actions*. London: Routledge & Keegan Paul.
- McCann, H. (1986). Intrinsic Intentionality. *Theory and Decision*, 20, 247-73.
- McCann, H. (1989). Intending and Planning: A reply to Mele. *Philosophical Studies*, 55, 107-110.
- McCann, H. (1991). Settled Objectives and Rational Constraints. *American Philosophical Quarterly*, 28, 19-30.
- McDowell, J. (1978). Are Moral Requirements Categorical Imperatives?. *Proceeding Aristotelian Society*, suppl. Vol. 52, 13-29.
- McDowell, J. (1994). *Mind and World*. London-Cambridge, Mass.: Harvard University Press.
- Mele, A. (1989). Intention, Belief, and Intentional Action. *American Philosophical Quarterly*, 26, 19-30.
- Mele, A. (1992). *Springs of Actions*. New York: Oxford University Press.
- Mele, A. (2000). Deciding to Act. *Philosophical Studies*, 100, 81-108.
- Mele, A. y Moser, K. (1994). Intentional Action. *Noûs*, 28, 1, 39-68.
- Miller, G., Galanter, E. y Pribram, K. (1983). *Planes y estructura de la conducta*. Madrid: Debate.
- Mosterín, J. (1978). *Racionalidad y acción humana*. Madrid: Alianza.
- O'Shaughnessy, B. (1980). *The Will, Vol 1 y 2*. Cambridge: Cambridge University Press.
- Pollack, M. (1992). The uses of plans. *Artificial Intelligence*, 57, 43-68.
- Real Academia Española (1992). *Diccionario de la Lengua Española*. Madrid: Espasa-Calpe.
- Ryle, G. (1967). *El concepto de lo mental*. Buenos Aires: Paidós.
- Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Searle, J. (1990). Collective Intentions and Actions. En P. Cohen, J. Morgan, M. Pollack. (eds.), *Intentions in Communication*, Cambridge, Mass.: MIT Press.
- Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, Mass.: MIT Press.
- Searle, J. (1995). *The Construction of Social Reality*. New York: The Free Press.
- Searle, J. (1997). *La construcción social de la realidad*. Barcelona-Buenos Aires-México: Paidós.
- Searle, J. (1998). *Mind, Language, and Society*. New York: Basic Books.
- Searle, J. (2000). *Razones para actuar. Una teoría del libre albedrío*. Oviedo: Ediciones Nobel.
- Schank, R. y R. Abelson (1977). *Scripts, Plans, Goals and Understanding*. Hillsdale: N.J., Lawrence Erlbaum Associates.
- Tomás de Aquino (1988). *Suma de Teología II*. Madrid: Biblioteca de Estudios Cristianos.
- Velleman, D. (1989). *Practical Reflection*. Princeton, N.J.: Princeton University Press.
- Velleman, D. (1992). What Happens When Someone Acts?. *Mind*, Vol. 101.403, July 1992, 461-481.
- Velleman, D. (1997). How To Share An Intention. *Philosophy and Phenomenological Research*, LVII, No 1, March 1997, 29-50.

Vermazen, B. (1993). Objects of Intention. *Philosophical Review*, 71, 223-65.

Wilson, G. (1989). *The Intentionality of Human Action*. Stanford: Stanford University Press.

Reseña: *Simulation and Similarity* de Michael Weisberg

MICHAEL WEISBERG

Simulation and Similarity: Using Models to Understand the World. Oxford: Oxford University Press, 2013, 208 pp.

ISBN: 978-0199933662

Itatí Branca¹

Recibido: 30 de agosto 2016.

Aceptado: 2 de septiembre de 2016.

La utilización de modelos para representar y comprender distintos fenómenos del mundo es una estrategia tan antigua como la ciencia misma. En el año 2000 AdC Egipcios y Babilónicos ya empleaban modelos matemáticos para intervenir e introducir mejoras en su vida cotidiana (Schichl, 2004). No obstante, el uso sistemático de estos dispositivos en ciencia puede enmarcarse a partir del siglo XIX, tratándose principalmente de modelos mecánicos (Bailer-Jones, 2013), y sólo recientemente el estudio de este tema cobró particular importancia en el marco de la filosofía de la ciencia y la epistemología.

Este interés epistemológico sobre la construcción de modelos y sus usos en las prácticas científicas comenzó a manifestarse tímidamente en 1950, a partir de la apreciación de la ubicuidad de su uso en ciencia, siendo analizado su rol en relación al desarrollo de las teorías (Bailer-Jones, 1999). Posteriormente, fue señalado cada vez con mayor énfasis el lugar privilegiado que ocupan los modelos en la ciencia, llegando a considerarse cierta autonomía en la actividad del modelado respecto de las teorías (Morgan & Morrison, 1999), o incluso a afirmarse que eran los modelos y no las teorías los auténticos vehículos del conocimiento científico (Suppe, 2000).

Ahora bien, se agrupan bajo la noción de “modelo” a una gran variedad de dispositivos, abstractos o concretos, tales como sistemas de ecuaciones matemáticas, gráficos, estructuras computacionales, físicas, entre otros. A su vez, estos son utilizados con fines muy variados como el de precisar predicciones, resumir datos, el diseño heurístico de experimentos, la representación y explicación de fenómenos, etc (Bogen, 2005). Esta heterogeneidad en las formas, contenidos y usos de los modelos, han convertido a su abordaje una tarea particularmente compleja.

El libro “Simulación y similaridad” de Michael Weisberg representa una contribución muy relevante al poner a disposición de científicos y filósofos un marco general, unificado, que aborda preguntas centrales que atraviesan la cuestión del modelado en ciencia: ¿Qué son los modelos?, ¿Qué tipos hay?, ¿Cómo se construyen?, ¿En qué consisten las idealizaciones?, ¿Cuál es el rol de las interpretaciones de los modeladores?, ¿Qué relaciones existen entre los modelos y los fenómenos del mundo?

Quizás una de las mayores virtudes del libro resida en el abordaje claro y organizado de todas estas cuestiones, siempre con la mirada puesta en las prácticas científicas y el desarrollo de ejemplos concretos. Este estilo da lugar a un libro sumamente accesible para principiantes y el público general, sin dejar de atender al mismo tiempo a discusiones claves, y plantear aportes novedosos que lo vuelven igualmente interesante

¹ CONICET/CIFyH – UNC

✉ itatibranca@gmail.com

Branca, Itatí; (2016); Reseña: *Simulation and Similarity*; *Epistemología e Historia de la Ciencia*; 1(1); 87-93. ISSN: 2525-1198



para cualquier experto en el área. En lo siguiente presentaremos brevemente algunos de estos desarrollos, para luego detenernos en el análisis de algunas cuestiones o aspectos que merecen ser profundizadas.

Continuando con la perspectiva pragmática de Giere (2004), Weisberg define a los modelos como “estructuras interpretadas que pueden ser utilizadas para representar un fenómeno real o imaginado” (Weisberg, 2013, p. 15). Estos dos componentes, la “estructura” y la “interpretación”, conforman el núcleo esencial sobre los que se erigen la mayoría de los aportes subsiguientes en el libro.

La consideración por los distintos tipos de estructuras constitutivas, lleva a Weisberg a distinguir entre tres clases de modelos utilizados en las prácticas científicas contemporáneas: *modelos concretos*; *modelos matemáticos*; y *modelos computacionales*. De cada una de los cuales desarrolla un ejemplo paradigmático al que recurrirá en lo sucesivo (el modelo hidráulico de la bahía de San Francisco; el modelo Lokta-Volterra de las relaciones presa-predador; y el modelo de segregación urbana de Schelling).

Presentada esta taxonomía, Weisberg se introduce de lleno en el análisis de la constitución o “anatomía” de los modelos. De acuerdo a esta propuesta, los modelos se componen por tres elementos claves: una estructura –concreta, matemática o computacional–, una interpretación² y una descripción. La descripción puede ser concreta o abstracta, y su función consiste en especificar al modelo (al menos parcialmente). La interpretación, a su vez, introduce un aspecto de especial relevancia en la propuesta de Weisberg: la intención de los modeladores. Dicha interpretación agrupa la definición de la *asignación*, esto es, cómo se mapean las distintas partes del sistema de interés (“target”) en el modelo; el *alcance* deseado por parte de los modeladores, qué aspectos del sistema se intentará representar mediante el modelo; y los *criterios de fidelidad* que capturan el grado de ajuste modelo-mundo aceptable para los modeladores. Esto incluye tanto a la exactitud de los resultados o comportamientos que arroja el modelo con respecto a los observados en el sistema de interés (*fidelidad dinámica*), como a la medida en que la estructura del modelo mapea la estructura del sistema en el mundo (*fidelidad representacional*).

A esta estructura general, Weisberg suma un factor que cumple un rol adicional en la construcción de los modelos: las ontologías populares. Estas involucran a las imágenes mentales, y la capacidad de apelar a escenarios ficcionales que tienen un lugar privilegiado a la hora de pensar, construir y analizar los modelos. Mediante la introducción de esta noción el autor intenta evadir a las perspectivas ficcionales de modelado. Perspectivas que además objeta mediante una serie de argumentos en los que no nos detendremos aquí pero que resultan convincentes.

En el libro, Weisberg también realiza un gran trabajo en conducir al lector a una visita guiada a la cocina del modelado en las prácticas científicas. Aquí presenta nuevamente dos grandes variedades de estilos, los modelos construidos con un sistema de interés específico como objeto, y modelos no orientados a un sistema concreto. En relación a estos tipos de modelos desarrolla múltiples ejemplos de las formas en que son construidos y analizados, señalando distintas particularidades que estos presentan en cada caso. En esta gran cocina, los objetivos y alcance esperado por los modeladores son la sal del modelado, los cuales se encuentran determinados a un nivel más general por los ideales representacionales adoptados. En el capítulo seis, Weisberg detalla estos distintos tipos de ideales que suelen regir la construcción, análisis y evaluación de modelos: *completitud*, *simplicidad*, *causal-1*, *maximización* (“maxout”) y *generalidad-p*.

² El término que Weisberg utiliza para hacer referencia a la interpretación es el de “Construal” que no tiene traducción directa al castellano, por lo que adoptaremos la palabra “interpretación” como equivalente.

A su vez, la adopción de algunos de estos ideales también determina el tipo de idealización que será implementada a la hora de introducir distorsiones en los modelos. Estas distorsiones pueden ser introducidas a fines de simplificar y reducir detalles para hacer el modelo más abordable en un principio siendo luego completado de manera progresiva (*idealizaciones galileanas*); focalizar el núcleo de relaciones causales más relevantes presentes en el sistema de interés (*idealizaciones mínimas*); o abordar un fenómeno de interés complejo mediante distintos modelos simples que varían en las características que representan del sistema y sus asunciones (*idealización de modelos múltiples*).

La perspectiva general de Weisberg sobre los modelos se completa luego en el capítulo ocho, en donde el mismo presenta su perspectiva de *similaridad* como la relación que debe establecerse modelo-mundo. Esta noción no involucra una forma prefijada y única en la que el modelo puede relacionarse al mundo sino que existen muchas relaciones posibles. Los objetivos teóricos de los científicos, su reconstrucción del sistema de interés y el contexto, son los que determinan en cada caso cuáles son los aspectos relevantes del sistema a ser representados y en qué grado los modelos deben ser similares en estos aspectos para considerarse adecuados (nótese aquí la fuerte influencia de los desarrollos de Giere, 1988). Con esta definición, el autor consigue dar respuesta a discusiones clásicas sobre la noción de similaridad, y a la crítica de que “cualquier cosa podría ser similar a otra”: El modelo debe ser similar al sistema en aspectos y grados específicos que son definidos por los modeladores en cada caso de acuerdo a sus intereses y teorías. Al mismo tiempo, la relación de similaridad no siempre se establece entre el modelo y el sistema de interés de manera directa. Frecuentemente, el modelo es contrastado con una representación matemática o abstracta de dicho sistema. De esta forma, el autor también consigue dar respuesta a algunos problemas clásicos en relación a la capacidad de representación de los modelos matemáticos, computacionales o abstractos, sobre fenómenos concretos/físicos.

Ahora bien, el punto más novedoso en la propuesta de Weisberg reside en el desarrollo de una sofisticada fórmula que permite evaluar el grado de similaridad entre el modelo y el sistema de interés. Esta fórmula se inspira en los trabajos psicológicos previos de Tversky (1977), quien estudió los juicios de similaridad o disimilaridad efectuados por sujetos experimentales. De acuerdo a estos desarrollos, los juicios de similaridad dependen en gran medida del peso que le otorgan los sujetos a las características de los objetos y cuáles consideran más relevantes para ser contrastadas. Esta valoración es formalizada –mediante la teoría de conjuntos– en una función que permite sopesar la relevancia de las distintas características y el grado en que son compartidas por ambos objetos. Mientras las características compartidas sean mayores a las no compartidas podrá decirse que estos son similares. Esta perspectiva le brindó a Weisberg un marco formal para introducir la noción de la “interpretación” de los modeladores. Del mismo modo en que los sujetos ponderan las características y definen a unas como más relevantes que otras, los modeladores otorgan un peso diferente a las características del sistema de acuerdo a sus objetivos, criterios de fidelidad y el contexto en el que se desarrolla el modelo. Sólo habiendo definido la relevancia de cada característica, puede valorarse luego la medida en que son representadas por el modelo. La ponderación del conjunto total de características, permite establecer el grado de similaridad modelo-sistema.

Por último, Weisberg presenta una herramienta epistémica más para el abordaje de los modelos: el análisis de robustez. Este tipo de análisis consiste básicamente en introducir variaciones en el modelo ya sea en los valores de los parámetros, los componentes vinculados a los mecanismos estructurales o incluso el modelo completo

modificando la forma en la que este es representado. Estas variaciones permiten determinar cuán robustos son los atributos mecánicos del modelo, y este en su totalidad. Weisberg defiende que esta prueba tiene un valor epistémico fundamental al permitirnos determinar la confiabilidad del modelo, como así también de las predicciones y explicaciones que obtenemos de él.

Habiendo realizado un recorrido por los contenidos abordados por Weisberg puede apreciarse ahora lo ambicioso que es el libro, y lo sustancial que resulta como una guía para el abordaje epistemológico de distintas cuestiones que atañen al tema del modelado en las prácticas científicas. Esta gran variedad de temas abordados en los distintos capítulos, sin embargo, se presenta a veces poco articulada en una estructura más general, y es tarea del lector ir entretejiendo estos contenidos. Quizás esto se deba a que Weisberg reúne en este libro muchos contenidos trabajados en publicaciones anteriores (Weisberg, 2006, 2007a, 2007b, 2012). Al mismo tiempo, como un buen libro de tal envergadura, deja abierta numerosas preguntas y puntos a profundizar. En esta dirección se nos presentan algunos temas que merecen una mayor atención.

El primer punto del que cualquier lector debiera quedar advertido es que este no es un libro sobre simulaciones. Más allá de la promesa en el título, Weisberg casi no se detiene en el tema a lo largo del libro, que como podrá notarse es sobre los modelos en ciencia. Si bien estas nociones, modelado y simulaciones, se encuentran estrechamente vinculadas (sobre todo en el caso de los modelos computacionales) existe una clara distinción entre las mismas (Maria, 1997). De hecho Weisberg las distingue, al hablar sobre análisis numérico menciona

Otras veces, el análisis numérico involucra simulación, computar el comportamiento del modelo usando un set de condiciones iniciales, lo que da un resultado sólo para ese set de condiciones iniciales (Weisberg, 2013, p.82, mi traducción).

Notando esta distinción que el mismo autor establece, no se explica luego el título elegido, o la falta de desarrollo sobre este tema específico.

En segundo lugar, atendiendo a un tema de importancia menor, podemos notar que los criterios que Weisberg utiliza para establecer su taxonomía de modelos no resultan claros. Como anticipamos, el autor enuncia que de acuerdo a la “estructura” componente pueden distinguirse modelos concretos, matemáticos o computacionales. Sin embargo, luego agrega que los modelos computacionales bien podrían ser descriptos como modelos matemáticos. Para Weisberg las razones para mantener la distinción entre estos modelos entonces residen en que son usados de distintas formas para “dar explicaciones científicas” (p.20). Mientras los modelos matemáticos centran su fuerza en los estados del sistema y relaciones entre variables, los modelos computacionales atienden a las reglas de transición o procedimientos, “Entonces mientras que ontológicamente no son distintos tipos de modelos, estos funcionan diferente en la práctica y tienen capacidades representacionales diferentes, por lo que deben serles dadas sus propias categorías” (Weisberg, p.20, mi traducción).

Si bien la distinción que Weisberg señala entre modelos matemáticos y computacionales es plausible, el criterio de distinción ahora ya no está puesto en las diferencias en la estructura, sino en la función y formas de uso. Con este criterio en mente, ahora se vuelve confuso clasificar en un mismo grupo de modelos concretos, a modelos como los construidos a escala y modelos de organismo. Si bien ambos tipos de modelos tienen una estructura física, estos también funcionan diferente en la práctica y tienen capacidades representacionales diferentes. Como bien señala Eric Winsberg (2013) en su propia reseña de este libro, lo que justifica las inferencias que podemos extraer es muy

distinto en el caso de los modelos de organismo y los modelos construidos. Mientras que en el primer caso nuestras inferencias se ven justificadas por una historia evolutiva o factores biológicos que el modelo organismo comparte con el sistema de interés; en el segundo caso estas se justifican por haber podido reconstruir a escala de una forma exacta las características del fenómeno de interés. En este sentido, los modelos construidos pueden ser erróneos, mientras que los modelos de organismo no. Estas, entre otras diferencias, nos obligarían a establecer mayores distinciones entre los modelos “concretos”, nótese que estos además incluyen a los modelos de experimentos y modelos de poblaciones.

Exigir una profundización exhaustiva que considere las características particulares de todas las clases de modelos en el desarrollo de una taxonomía es un contrasentido, no obstante, sí es deseable que las distinciones generales trazadas conserven un criterio claro.

Otra cuestión llamativa del libro es que al introducirse en el tema de la idealización Weisberg omite enmarcarla en la distinción clásica de abstracción e idealización desarrollada en la literatura (Cartwright, 1999; Harré, 2002; Jones, 2005). Para Weisberg, todos los casos presentados son idealizaciones, las cuales consisten en introducir distorsiones en el modelo de manera intencional. Estas “distorsiones” involucran la simplificación, restringir detalles, asumir características falsas no relevantes, entre otros. La única mención que realiza del término abstracción, es en el marco de los modelos mínimos para referir a la exclusión de detalles no causales que estos presentan. Sin embargo, en los otros tipos de “idealización” también se implementa la abstracción (ya sea para “simplificar” en el caso de las *idealizaciones galileanas*, o para focalizar sólo alguna característica en el caso de los *modelos múltiples*). Distinguir las nociones de abstracción e idealización permitiría esclarecer el modo en que estas estrategias son utilizadas en los distintos casos presentados por Weisberg. Asimismo, hubiera resultado interesante que el autor realizara algún abordaje del tema de la introducción de elementos falsos o ficciones en los modelos, y los usos y funciones que este tipo de estrategia tiene en ciencia³.

Finalmente, existe una última cuestión a revisar que atañe a un eje central del libro: la perspectiva de similaridad basada en la ponderación de características. A pesar de la flexibilidad que tiene la fórmula desarrollada por Weisberg para ser aplicada a la gran variedad de modelos utilizados en las prácticas científicas, esta no resulta del todo satisfactoria para dar cuenta de la relación de similaridad. Es en todo caso una buena propuesta para evaluar la adecuación del modelo en relación a los objetivos e interpretaciones de los modeladores, pero no para determinar cuan adecuado es el ajuste modelo-mundo.

La definición de Weisberg de que un modelo debe ser similar al sistema en aspectos específicos que se determinan de acuerdo a los objetivos adoptados por los modeladores, es en principio una idea plausible. El problema en la generalidad de esta propuesta surge cuando se evalúa la fórmula que presenta Weisberg. Como anticipamos, al aplicar esta fórmula para determinar el grado de similaridad modelo-sistema de interés, se otorga un peso específico a cada característica a contrastar. Esta valoración de cada característica depende de los objetivos y criterios de fidelidad adoptados por el modelador. Así, por ejemplo, si el interés del modelador fue el de identificar mecanismos *posibles* que expliquen el comportamiento del sistema, no será valorado que el modelo presente un alto el criterio de fidelidad representacional respecto a los mecanismos. Esto se verá luego

³ Weisberg presenta en el capítulo siete a los modelos hipotéticos que son altamente ficcionales, sin embargo, esto es muy lejano a la introducción de elementos falsos o ficciones en modelos no hipotéticos como puede ser el uso de la noción de poblaciones infinitas.

reflejado en la fórmula, no siendo valorada la similaridad a nivel de mecanismos pero sí a nivel de los atributos. El problema entonces es que un modelo que presenta un “mecanismo posible” podría ser determinado altamente similar al sistema de interés, mientras que sólo presenta similaridad respecto a los resultados que arroja. Como puede apreciarse, esto muestra como en verdad la fórmula no nos permite apreciar el grado de similaridad del modelo al sistema, sino cuán adecuado es respecto a los ideales establecidos por los modeladores. Esta noción de similaridad, además, no nos dice nada acerca de cómo establecer si las relaciones representacionales son adecuadas, ni cómo establecerlas.

Más allá de todos estos señalamientos, el libro de Weisberg es realmente un mapa extenso que aborda numerosas cuestiones del modelado en las prácticas científicas, invitando al lector a introducirse en este complejo mundo. El libro tiene además la virtud de mostrar al lector una serie de problemas filosóficos que debieran ser abordados a la hora de pensar en los modelos y sus usos en las prácticas científicas.

1. Agradecimientos

La revisión conjunta y discusiones sobre este libro que llevamos a cabo con Juan Durán, Manuel Barrantes y Andrés Ilcic representaron un aporte fundamental a pensar los temas aquí presentados. A ellos quiero agradecer por ese espacio de discusión tan motivador que hemos construido.

2. Referencias

- Bailer-Jones, D. M. (1999). Tracing the Development of Models in the Philosophy of Science. En L. Magnani, N. J. Nersessian, & P. Thagard (Eds.), *Model-Based Reasoning in Scientific Discovery*; 23-40. Springer
- Bailer-Jones, D. M. (2013). *Scientific Models in Philosophy of Science*. Pittsburgh, Pa.: University of Pittsburgh Press.
- Bogen, J. (2005). Regularities and causality; generalizations and causal explanations. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 397-420.
- Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge, UK: New York, NY: Cambridge University Press.
- Giere, R. N. (2004). How Models Are Used to Represent Reality. *Philosophy of Science*, 71(5), 742–752.
- Harre, R. (2002). *Cognitive Science: A Philosophical Introduction*. SAGE Publications.
- Jones, M. R. (2005). Idealization and Abstraction: A Framework. En M. R. Jones & N. Cartwright (Eds.), *Correcting the Model: Idealization and Abstraction in the Sciences*; 173–218. Rodopi.

- Maria, A. (1997). Introduction to Modeling and Simulation. En *Proceedings of the 29th Conference on Winter Simulation*; 7–13. Washington, DC, USA: IEEE Computer Society.
- Morgan, M. S., & Morrison, M. (1999). *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge University Press.
- Schichl, H. (2004). Models and the History of Modeling. En J. Kallrath (Ed.), *Modeling Languages in Mathematical Optimization*. pp. 25-36. Springer US.
- Suppe, F. (2000). Understanding Scientific Theories: An Assessment of Developments, 1969-1998. *Philosophy of Science*, 67(3), 115.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327-352.
- Weisberg, M. (2006). Robustness Analysis. *Philosophy of Science*, 73(5), 730-742.
- Weisberg, M. (2007a). Three Kinds of Idealization. *Journal of Philosophy*, 58, 207–233.
- Weisberg, M. (2007b). Who is a Modeler? *The British Journal for the Philosophy of Science*, 58(2), 207-233. Weisberg, M. (2012). Getting Serious about Similarity. *Philosophy of Science*, 79(5), 785-794.
- Winsberg, E. (2013). [Review of *Review of «Simulation and Similarity: Using Models to Understand the World»*, por M. Weisberg]. Recuperado a partir de <http://ndpr.nd.edu/news/38637-simulation-and-similarity-using-models-to-understand-the-world/>