

## **Sobre el poder explicativo de los enunciados cognoscitivos**

*(On the explanatory power of cognitive statements)*

**José E. Burgos(\*)**

CEIC - Universidad de Guadalajara

El problema de caracterizar las explicaciones científicas ha recibido considerable atención en la filosofía de la ciencia. El tema posee múltiples aspectos, tanto epistémicos como metafísicos, que se relacionan, solapan y entretienen de muchas formas diferentes. Cuando se empieza una discusión epistemológica, resulta inevitable caer en una discusión metafísica, y viceversa, encontrándonos a lo largo del proceso con toda clase de consideraciones lingüísticas. Es imposible, entonces, discutir un aspecto sin al menos mencionar el otro, y hacer ambas cosas fuera de un contexto lingüístico, y el presente artículo no será una excepción. Sin embargo, aquí me concentraré sobre lo epistémico, aunque, por supuesto, no ignoraré por completo lo metafísico. Mi objetivo no es afirmar la suficiencia de lo epistémico respecto a lo metafísico y lo lingüístico, sino, más bien, delimitar la discusión.

En particular, mi objetivo principal es reflexionar críticamente acerca de la legitimidad y valor epistémico las explicaciones cognoscitivas. El lector podría pensar, con razón, que voy a presentar una crítica más del conductista radical típico, de las mismas que ya hemos escuchado hasta la saciedad. Sin embargo, mi crítica incluirá por igual a los conductistas radicales mismos y a los filósofos de la ciencia. El presente artículo, entonces, debe verse como una reflexión crítica acerca de lo que cognoscitivistas, filósofos de la ciencia, y conductistas radicales han dicho acerca de la legitimidad y valor epistémico de las explicaciones cognoscitivas.

Mi análisis consistirá de dos grandes partes. En una primera parte (la cual constituye el grueso del artículo), muestro a los conductistas radicales que las

(\*)Agradezco a Alejandro Tomasini Bassols y François Tonneau por comentarios esclarecedores a versiones previas de este manuscrito. Cualquier correspondencia sobre este artículo puede ser enviada al autor. Universidad de Guadalajara, Centro de Estudios e Investigaciones en Comportamiento, Av. 12 de Diciembre # 204, Col. Chapalita, Código Postal 45030, Guadalajara, Jal., (A.P. 5-374), MÉXICO. E-mail: jburgos@cucba.udg.mx. Sitio web: <http://udgserv.cenear.udg.mx/~ceip/>

explicaciones cognoscitivas son epistémicamente legítimas, al menos bajo tres modelos de la explicación, a saber, el modelo nomológico-deductivo (ND) de Hempel y Oppenheim (1948), el modelo de la relevancia estadística (RE) de Salmon (1984) y el modelo de la inferencia a la mejor explicación (IME), tal y como ha sido elaborado por Lipton (1991). Esta conclusión va en contra de lo que el conductista radical típico argüiría (i.e., que las explicaciones cognoscitivas son epistémicamente ilegítimas, puesto que apelan a estados y procesos inobservables) y, por tanto, favorece tanto a psicólogos cognoscitivos como a filósofos de la ciencia. Sin embargo, en la segunda parte critico a los dos últimos, arguyendo que las explicaciones provenientes de las *neurociencias* son tan epistémicamente legítimas como las explicaciones cognoscitivas, exactamente bajos esos tres mismos modelos. Ello plantea el problema de determinar cuáles explicaciones son las más *satisfactorias*, problema que supone una distinción entre *legitimidad* y *satisfacción* epistémica. Concluyo que las explicaciones neurocientíficas resultan indudablemente más satisfactorias que las explicaciones cognoscitivas, al menos bajo los modelos ND y RE. Bajo el modelo IME, las explicaciones cognoscitivas pueden ser más ‘satisfactorias’ que las neurocientíficas, pero en un sentido un tanto peculiar y muy discutible.

Mi objetivo podría descartarse de entrada, arguyéndose que la psicología cognoscitiva está en principio errada al buscar *explicaciones*, puesto que la psicología es (o debe ser) una disciplina estrictamente *descriptiva*. El problema con este tipo de consideración es que, en realidad, no queda claro si es, a su vez, descriptiva o prescriptiva. Si es descriptiva (i.e., si supone que la psicología *es* de hecho exclusivamente descriptiva), entonces es obviamente falsa. Si es prescriptiva, es decir, si ordena que la psicología *debe ser* exclusivamente descriptiva, entonces no queda claro exactamente por qué ello debe ser así. En general, una prescripción de este tipo se basa sobre alguna hipótesis justificada *a priori* sobre la base de consideraciones lógicas (e.g., regresión al infinito de preguntar ‘¿por qué?’) que son fácilmente refutables. Una variante común de esta línea de argumentación es considerar que los términos ‘explicación’ y ‘descripción’ son sinónimos. Pero, de nuevo, si tal consideración es descriptiva, entonces es falsa, puesto que tales términos son usados de modos muy diferentes, tanto en ciencia como en filosofía, y si es prescriptiva, entonces es arbitraria y, por tanto, injustificable a la luz de tales usos.

Mi objetivo podría también rechazarse en términos de una visión *instrumentalista* de las teorías, arguyéndose que éstas *no son más* que meros instrumentos de predicción, *sin poder explicativo alguno*. Sin embargo, el psicólogo cognoscitivo debe ser oficialmente realista (pero ver Tonneau y Tamayo, este número), en el sentido de que debe considerar sus enunciados como literalmente verdaderos (aunque el más mínimo intento de criticar esta posición lo empuje hacia un instrumentalismo). De lo contrario,

pierde el derecho de afirmar que sus inferencias son verdaderas, e incluso de *justificar* su investigación en términos de la verificación (o falsación) de las mismas. Además, si el científico (incluyendo al psicólogo) cognoscitivo en general pretende *explicar* la conducta (e.g., Simon, 1992; cf. Zentall, 2001), entonces debe necesariamente ser realista, puesto que, de nuevo, bajo el instrumentalismo, nuestras teorías no deben ser usadas como recursos explicativos. En el presente artículo, entonces, supondré que ‘explicación’ y ‘descripción’ no son sinónimos (que la psicología de hecho pretende explicar, además de describir), que la regresión al infinito disparada por preguntas del tipo ‘por-qué’ puede ser detenida en cualquier punto arbitrario sin mayor problema, y que los enunciados y que los psicólogos cognoscitivos son en su mayoría realistas (Amundson, 1985; pero ver Tonneau y Tamayo, este número), en el sentido de que sus teorías tienden a ser más que meros instrumentos de predicción (i.e., tienden a ser consideradas como potencialmente verdaderas, en el sentido de que pueden referirse a una realidad que existe objetivamente).

Los tres modelos de la explicación que he escogido difieren en aspectos críticos. Por una parte, los modelos ND y RE se ocupan exclusivamente del problema de la *justificación* (el problema de qué es una explicación adecuada, correcta, científica, o de cómo evaluamos las explicaciones como productos de nuestros actos de explicar). Sobre esta base, tales modelos podrían calificarse como *evaluativos* o *justificativos*. El modelo IME, por su parte, se ocupa principalmente del problema de la *descripción* de nuestras prácticas de inferencia, como punto de partida hacia una solución al problema epistemológico de la justificación.

Por otra parte, ND y RE se asemejan en proponer que toda explicación adecuada (de hechos particulares) involucra *al menos* una relación de *subsunción* de (una proposición que expresa la ocurrencia de) un hecho particular bajo ciertas *leyes*. En ND, tales leyes son deterministas, mientras que en RE, son estadísticas o probabilísticas (aunque el segundo puede también incorporar leyes deterministas, asignando el valor de 1 a las leyes probabilísticas). Sin embargo, en el modelo ND, la relación de subsunción es una relación de inferencia (deductiva o inductiva, en caso de que las leyes sean deterministas o estadísticas, respectivamente), mientras que en el modelo RE tal relación es de pertenencia de los individuos involucrados en el hecho de interés a una cierta clase que constituye una partición estadísticamente relevante del dominio de interés (ver más adelante).

Me serviré de dos distinciones estándar en la literatura filosófica sobre la explicación científica. La primera es aquella entre *explanandum* y *explanans*. El *explanandum* es el fenómeno o hecho a ser explicado, el cual es lingüísticamente expresado mediante un enunciado denominado ‘*explicandum*’. Por su parte, el *explanans* es aquello que se ofrece para dar cuenta del *explanandum*, y es lingüísticamente expresado

mediante un conjunto de enunciados que constituyen el *explicans*. La segunda distinción es aquella entre enunciados generales y enunciados particulares. Un enunciado general es aquel que se refiere a múltiples individuos a la vez, mientras que un enunciado particular se refiere a un individuo concreto, espaciotemporalmente localizado. Esta distinción resulta crítica a la hora de separar entre dos tipos de explicaciones, a saber, aquellas cuyo *explicandum* es un enunciado particular versus aquellas cuyo *explicandum* es un enunciado general. La distinción es también útil para analizar el *explicans* en dos tipos de enunciados, a saber, *leyes* y *condiciones iniciales*.

Como punto de partida, me concentraré sobre explicaciones cuyo *explicandum* es un *enunciado particular*, es decir, un enunciado sobre un individuo concreto (espaciotemporalmente localizado) que ha sido sometido a cierto tratamiento y ha mostrado ciertas modificaciones como resultado del mismo. Se podría argüir que tales explicaciones no califican como científicas, en tanto en cuanto lo que (supuestamente) le interesa a la ciencia, en última instancia, es dar cuenta de *regularidades* y, por tanto, alcanzar explicaciones cuyos *explicanda* sean enunciados generales (o leyes). En esto estoy completamente de acuerdo. Sin embargo, las explicaciones de hechos particulares poseen una relevancia práctica que resulta difícil de ignorar en la psicología. Además, las explicaciones de eventos particulares son mucho más fáciles de caracterizar que las explicaciones de regularidades, ya que las últimas se ubican en esa frontera difusa y por demás problemática entre explicación y teoría. Ciertamente no evitaré esta frontera, ya que hacia el final del artículo le daré cierta consideración. Pero resulta conveniente empezar por lo más sencillo.

Utilizaré como ejemplo a lo largo de mi análisis uno de los fenómenos más básicos, mejor establecidos y estudiados en la psicología, a saber, el *condicionamiento pavloviano* (o 'clásico' o 'respondiente'), en su forma más simple. Este fenómeno es bien conocido, tanto dentro como fuera de la psicología. Supóngase que un organismo particular ha sido expuesto a un tratamiento en el cual instancias de un cierto tipo de estímulo (denominado 'condicional', o EC) fueron presentadas en una cierta relación de contigüidad espacial y temporal con instancias de otro tipo de estímulo ('incondicional', o EI) que originalmente evocaba o producía una fuerte respuesta. Este es el procedimiento estándar de condicionamiento pavloviano y su efecto típico es que el EC eventualmente llega a evocar el mismo tipo de respuesta que antes del tratamiento era evocada sólo por el EI. Supongamos, entonces, que dicho efecto fue observado, ante lo cual podemos preguntar '¿por qué?'.

---

## LA LEGITIMIDAD EPISTÉMICA DE LAS EXPLICACIONES COGNOSCITIVAS

### *Aplicación del modelo ND*

Responder esa pregunta, en términos del modelo ND, resulta en un *argumento deductivo*. La forma más simple posible de tal argumento sería la siguiente:

Todo organismo tratado con un procedimiento estándar de condicionamiento pavloviano, incrementa su respuesta ante el EC.

Este organismo fue tratado con un procedimiento estándar de condicionamiento pavloviano.

Por tanto,

Este organismo incrementó su respuesta ante el EC.

Si formalizamos este argumento en el lenguaje de la lógica de predicados, obtenemos lo siguiente:

1.	$(\forall x)[(Ox \wedge Px) \rightarrow Rx]$	Ley
2.	$Oa \wedge Pa$	Condiciones iniciales
3.	$(Oa \wedge Pa) \rightarrow Ra$	EU 1
4.	$Ra$	PP 2, 3.

Las letras **O**, **P** y **R** denotan, respectivamente, los predicados ‘es un organismo’, ‘fue sometido a un procedimiento de condicionamiento pavloviano’ e ‘incrementó su repuesta ante el EC’. En este último predicado, se supone que la respuesta es una reacción biológicamente específica a cierto tipo de estímulo (el EI) y que, antes del tratamiento, no era evocada por el EC<sup>1</sup>. Por su parte, los símbolos ‘ $\forall$ ’, ‘ $\wedge$ ’, ‘ $\rightarrow$ ’ denotan, respectivamente, ‘para todo’, ‘y’ y ‘si... entonces...’, mientras que  $x$  denota un organismo cualquiera y  $a$  denota un organismo particular. Las abreviaciones EU y PP denotan las reglas de inferencia deductiva conocidas respectivamente como ‘eliminación del

<sup>1</sup> El argumento podría incluir estos y muchos otros detalles, pero prefiero mantenerlo lo más simple posible.

cuantificador universal' y 'modus ponendo ponens'. Los números justo a la derecha de cada abreviación denotan las líneas del argumento a las cuales fueron aplicadas dichas reglas para obtener la línea en cuestión. Así, por ejemplo, la Línea 3 fue obtenida aplicando la regla de eliminación del cuantificador universal al enunciado que constituye la Línea 1. Tal aplicación, como el nombre de la regla lo indica, involucra eliminar el símbolo ' $\forall$ ' del enunciado, junto con la  $x$  que está justo a su derecha, y reemplazar las demás  $x$ s por una constante individual que denote un organismo en particular. De este modo, lo que se dice de *todos* los organismos en el enunciado de la Línea 1, en el enunciado de la Línea 3 se dice de un organismo en particular.

He obtenido así un argumento deductivamente válido con dos premisas (Enunciados 1 y 2) que constituyen el *explicans* y una conclusión que expresa el hecho a ser explicado, la cual constituye el *explicandum*. De las dos premisas, el Enunciado 1 expresa una *ley empírica* y el Enunciado 2 expresa las condiciones iniciales. El argumento resultante representa una explicación de un hecho particular, la cual satisface las condiciones de adecuación (tanto formales como materiales) que son estipuladas por el modelo ND. Materialmente, se supone que tanto las premisas como la conclusión son enunciados *empíricamente verdaderos*. Formalmente, la conclusión es una consecuencia lógica de las premisas. La explicación en cuestión pareciera ser completa, en el sentido de no requerir premisa adicional alguna, por lo cual parecería innecesaria la postulación de *mecanismos subyacentes*. Antes de elaborar esta consideración, sin embargo, debo hacer tres salvedades.

Primero, tenemos el problema de la *justificación* de la ley misma. Usualmente, diríamos que dicha ley expresa una '*generalización inductiva*' a partir de una regularidad empírica observada, en el sentido de que es un enunciado universal que se supone como verdadero a pesar de que sólo algunos individuos del dominio han sido observados. Pero ¿sobre qué bases hacemos tal inferencia? Este es el problema clásico o *humano* de la inducción, que se nos presenta inmediatamente, antes de siquiera plantearnos el problema de los mecanismos subyacentes del comportamiento, y que se presentará nuevamente (bajo una forma un tanto diferente) cuando tales mecanismos sean introducidos. Dicho problema se refiere a la imposibilidad de justificar, mediante la lógica o mediante la experiencia, la inducción como forma de inferencia. Aun cuando representa uno de los problemas centrales de la filosofía de la ciencia, no voy a abordarlo aquí.

Segundo, las limitaciones del modelo ND son bien conocidas. Una de ellas es que es demasiado flexible, por cuanto no restringe qué califica como regularidad nómica y qué como regularidad accidental, lo cual, a su vez, permite la construcción de argumentos que, al menos intuitivamente, no parecen ser explicaciones legítimas. Aun cuando éste también es un problema importante, los psicólogos del aprendizaje están de acuerdo en

que la regularidad que he escogido no es accidental, aunque no me ocuparé del sentido exacto en el cual constituye una regularidad nómica, por cuanto ello nos llevaría a consideraciones metafísicas acerca de la causalidad que me desviarían considerablemente de mi objetivo principal.

Tercero, el modelo ND no nos permite caracterizar explicaciones cuyos *explicans* incluyen regularidades estadísticas. Este problema es particularmente agudo en el tipo de regularidad que he escogido, ya que se sabe que no todos los organismos aprenden a evocar la respuesta en presencia del EC. Esto plantea la cuestión de si el tipo de fenómenos que estudiamos en la psicología son fundamental, irreduciblemente *probabilísticos* o si, por el contrario, las regularidades estadísticas son simplemente el resultado de nuestra ignorancia y pueden ser en principio reducidas a regularidades deterministas. Tampoco voy a discutir este problema en el presente artículo, así como tampoco las limitaciones de la solución propuesta por Hempel al problema de construir explicaciones cuyos *explicans* incluyen leyes estadísticas (el modelo estadístico-inductivo, con su requisito de máxima especificidad). Sólo diré que *si* suponemos que las leyes en psicología son fundamentalmente probabilísticas y que una ley debe expresar una regularidad nómica, entonces debemos abandonar el modelo ND y, con ello, la idea de que las explicaciones son argumentos (deductivos o inductivos).

### *Aplicación del modelo RE*

Una alternativa sería utilizar el modelo RE, en términos del cual podemos expresar simbólicamente el tipo de explicación que he venido considerando de la siguiente manera:

1.  $O = \{x|Ox\}$  (clase inicial de referencia).
2.  $\pi(Ox|Rx) = p_1$  (probabilidad inicial de  $R$  en  $O$ ).
3.  $O_1 = \{x|Ox \wedge Px\}$ ,  $O_2 = \{x|Ox \wedge Cx\}$  (partición de  $O$  en  $O_1$  y  $O_2$ ).
4.  $\pi(Ox|Rx \wedge Px) = p_2$  (probabilidad posterior, luego de la partición).
5.  $p_2 \gg p_1$  (la partición de  $O$  en  $O_1$  y  $O_2$  es estadísticamente relevante).
6.  $a \in O_1$  (el organismo  $a$  pertenece a  $O_1$ ).

La Línea 1 de esta explicación es la definición de la clase inicial de referencia, la cual especifica el 'dominio', 'universo' o el tipo de individuos de interés al cual se aplica la explicación en cuestión. En este caso, he elegido la clase  $O$  de los organismos como la clase inicial de referencia, la cual viene definida por el predicado 'es un organismo'. La Línea 2 especifica la probabilidad de que un organismo cualquiera en condiciones naturales (es decir, sin haber sido sometido a un procedimiento pavloviano) reaccione de cierta manera ante cierto tipo de estímulo (en particular, de que reaccione

ante un estímulo biológicamente no específico, de la misma manera en la que reacciona a un estímulo biológicamente específico). La Línea 3 involucra una *partición* de **O** en dos subclases, a saber, la clase de organismos sometidos a un procedimiento estándar de condicionamiento pavloviano ( $O_1$ ) y la clase de organismos sometidos a algún procedimiento de control ( $O_2$ ), el cual puede ser o un procedimiento verdaderamente aleatorio o uno explícitamente desapareado (ver Rescorla, 1967). Esta segunda clase es definida en términos del predicado **C**, que denota 'fue sometido a un procedimiento de control'.

El núcleo del modelo RE se encuentra en la Línea 4, en la cual se compara la probabilidad anterior con la probabilidad posterior a la partición. Si las dos probabilidades difieren considerablemente (sin importar en qué dirección o la magnitud absoluta de cada probabilidad), entonces se considera que la partición en cuestión es *estadísticamente relevante*. En el tipo de fenómeno que nos ocupa, sabemos que la probabilidad de que un organismo responda ante un EC, luego de haber sido sometido a un procedimiento estándar de condicionamiento pavloviano, es considerablemente mayor que la probabilidad de que responda ante ese mismo EC luego de haber sido sometido a algún procedimiento de control. Bajo el modelo RE, entonces, se diría que el organismo tratado (designado como *a*) respondió ante el EC porque pertenece a la clase de los organismos que han sido tratados con un procedimiento de condicionamiento clásico, y la probabilidad de que ello ocurra en esta clase es sustancialmente mayor que la probabilidad de que ocurra en condiciones naturales o luego de un procedimiento de control<sup>2</sup>.

Al igual que la explicación tipo ND, la explicación RE de arriba pareciera no requerir de suposición adicional alguna. En este sentido, la aplicación de dos modelos justificativos de la explicación pareciera hacer innecesaria la introducción de términos de corto cognoscitivo. Ello pareciera favorecer el rechazo que el conductista radical típico muestra hacia las explicaciones cognoscitivas. De hecho, las dos explicaciones que he construido hasta ahora representan precisamente el tipo de explicación que es preferida por el analista conductual típico, a saber, explicaciones que apelan a leyes empíricas que hayan sido derivadas mediante un análisis experimental de la conducta (ampliamente concebido). Sin embargo, para ser justos, un psicólogo cognoscitivo podría objetar que el término 'organismo' (denotado por el predicado '**O**') en las dos

<sup>2</sup> La relevancia estadística de ciertas particiones, en sí misma, puede no tener valor explicativo alguno, en el sentido de que puede deberse a una mera correlación estadística, sin conexión *causal* alguna entre los factores de interés. Para solucionar esto, Salmon (1998) aclaró que la relevancia estadística era una condición necesaria más no suficiente para que una partición posea valor explicativo. Salmon propuso que las particiones estadísticamente relevantes deben ir acompañadas por *leyes causales*. Más adelante examinaré esta modificación que el mismo Salmon le hizo a su propio modelo.



explicaciones anteriores es *primitivo*, es decir, su significado no ha sido analizado. Cuando se lleva a cabo un análisis de ese término, continúa esta objeción, se producen nuevos predicados que nos permiten ampliar nuestras explicaciones, haciéndolas más informativas y, en esa medida, más completas y heurísticas. A continuación, examino esta posibilidad.

### *Analizando el predicado 'O'*

La manera usual de analizar el significado del predicado 'O' cognoscitivamente consiste en conceptualizar al organismo como un *sistema cognoscitivo*, es decir, como un conjunto de eventos y procesos que supuestamente ocurren 'dentro' del organismo cuando éste se comporta, juegan un papel causal en la conducta y no son interpretados ni conductual ni biológicamente. En el fenómeno que nos ocupa, un ejemplo de análisis cognoscitivo de O lo encontramos en la teoría de Rescorla y Wagner (1972), o RW en adelante. Según la *interpretación cognoscitiva* de esta teoría, ofrecida por los mismos autores,

“...los organismos aprenden sólo cuando los eventos violan sus expectativas. Los organismos se forman ciertas expectativas acerca de los eventos que ocurren después de un estímulo complejo; las expectativas originadas por ese complejo y sus estímulos componentes se modifican sólo cuando los eventos consecuentes difieren de la expectativa compuesta.” (p. 75)<sup>3</sup>.

Debo aclarar que el fenómeno básico de condicionamiento pavloviano pareciera involucrar un sólo EC. Sin embargo, el postulado central de la teoría RW es que las expectativas que un organismo se forma respecto a un cierto estímulo dependen de las expectativas que ya se ha formado respecto al *resto* los estímulos que han ocurrido y ocurren junto con el EC en cuestión. Para la teoría, entonces, toda situación de condicionamiento pavloviano involucra *estímulos compuestos* o *constelaciones de estímulos*. En el caso del arreglo típico, tenemos que toda ocurrencia del EC viene acompañada por un *contexto*, es decir, el conjunto de estímulos que constituyen el ambiente experimental y que permanecen relativamente estáticos a lo largo de una sesión experimental, y de una sesión a otra.

<sup>3</sup> Cabe destacar que Rescorla se encuentra entre los (pocos) psicólogos cognoscitivos que parecen ser verdaderamente instrumentalistas. Wagner, por su parte, quizás no sea tan instrumentalista como Rescorla, pero puede muy bien ser considerablemente menos realista que la mayoría de los psicólogos cognoscitivos. Por ello, al incluir esta cita no estoy afirmando que estos autores son de hecho realistas. Más bien, estoy asumiendo que la mayoría de los psicólogos cognoscitivos del aprendizaje que usan la teoría RW como guía para su investigación experimental en condicionamiento pavloviano toman esta interpretación de forma realista (es decir, como potencialmente verdadera, y, en esa medida, como poseyendo valor explicativo).

Según la teoría RW, entonces, el procedimiento estándar de condicionamiento pavloviano causa un incremento negativamente acelerado en la expectativa por parte del organismo de que el EI ocurrirá inmediatamente después del EC, lo cual, a su vez, causa un incremento en el responder del organismo ante el EC. Esta teoría puede ser vista como cognoscitiva en el sentido de que utiliza términos tales como 'expectativa', 'fuerza asociativa', y 'sorpresa', los cuales son definidos como procesos que, supuestamente, ocurren en un nivel de observación diferente del conductual (*viz.* 'dentro del organismo'), que no son biológicamente interpretados y que juegan un papel causal en la conducta. Por el momento, le daré el beneficio de la duda metafísica a la teoría y aceptaré que tales procesos poscen una existencia objetiva y que juegan un papel causal en la conducta.

### *Introduciendo enunciados cognoscitivos*

Cuando analizamos el término '**O**' cognoscitivamente, podemos generar otros enunciados, además de aquellos que ya teníamos, ampliando así las explicaciones que hemos obtenido hasta ahora. El problema es si tales ampliaciones constituyen explicaciones epistémicamente legítimas bajo los modelos ND y RE. Considérese el siguiente argumento, el cual se deriva de un análisis cognoscitivo del término 'organismo', y atiende a la conclusión de Derr y Thompson (1992) de que enunciados cognoscitivos deben incorporarse como leyes y no como condiciones iniciales (cf. Hempel y Oppenheim, 1948):

Todo organismo tratado con un procedimiento estándar de condicionamiento pavloviano incrementa su expectativa de que el EI ocurrirá luego del EC.

Todo organismo que incrementa su expectativa de que el EI ocurrirá luego del EC incrementa su respuesta ante el EC.

Este organismo fue tratado con un procedimiento estándar de condicionamiento pavloviano.

Por tanto,

Este organismo incrementó su respuesta ante el EC.

Formalizado en términos del modelo ND, obtenemos el siguiente argumento:

1.	$(\forall x)[(Ox \wedge Px) \rightarrow Ex]$	Ley RW1
2.	$(\forall x)[(Ox \wedge Ex) \rightarrow Rx]$	Ley RW2
3.	$Oa$	Condición inicial
4.	$Pa$	Condición inicial
5.	$(Oa \wedge Pa) \rightarrow Ea$	EU 1
6.	$Oa \wedge Pa$	Adjunción 3, 4
7.	$Ea$	PP 5, 6
8.	$(Oa \wedge Ea) \rightarrow Ra$	EU 2
9.	$Oa \wedge Ea$	Adjunción 3, 7
<hr/>		
10.	$Ra$	PP 8, 9.

Alternativamente, si consideramos que RW1 y RW2 son leyes fundamentalmente probabilísticas, podemos construir la siguiente explicación RE:

1.  $O = \{x|Ox\}$  (clase original de referencia).
2.  $O_1 = \{x|Ox \wedge Px\}$ ,  $O_2 = \{x|Ox \wedge Cx\}$  (Partición 1).
3.  $O_{1,1} = \{x|Ox \wedge Px \wedge Ex\}$ ,  $O_{1,2} = \{x|Ox \wedge Px \wedge \neg Ex\}$  (Partición 2).
4.  $\pi(Ox|Px \wedge Ex) = p_1$ ,  $\pi(Ox|Ox \wedge Px \wedge \neg Ex) = p_2$  (probabilidades posteriores a la Partición 2).
5.  $p_1 \gg p_2$  (la Partición 2 es estadísticamente relevante).
6.  $O_{1,1,1} = \{x|Ox \wedge Px \wedge Ex \wedge Rx\}$ ,  $O_{1,1,2} = \{x|Ox \wedge Px \wedge Ex \wedge \neg Rx\}$  (Partición 3).
7.  $\pi(Ox|Px \wedge Ex \wedge Rx) = p_3$ ,  $\pi(Ox|Ox \wedge Px \wedge Ex \wedge \neg Rx) = p_4$  (probabilidades posteriores a la Partición 3).
8.  $p_3 \gg p_4$  (La Partición 3 es estadísticamente relevante).
9.  $a \in O_{1,1,1}$  (pertenencia de  $a$  a  $O_{1,1,1}$ ).

Nótese que el ambiente (representado por el predicado  $P$ ) permanece como un aspecto integral, tanto como los (supuestos) cambios en la expectativa. Es imposible prescindir del ambiente en nuestras explicaciones de la conducta, aunque, como hemos visto, es posible prescindir de términos y enunciados cognoscitivos, si consideramos al predicado como  $O$  primitivo. Explicaciones *minimamente aceptables* de la conducta deben apelar al medio ambiente, aun cuando concedamos que pueden incompletas. Porque, si bien cambios observados en la conducta pueden deberse a cambios en las expectativas del organismo, debemos preguntarnos a qué se deben estos últimos, ante

lo cual, debemos responder, inevitablemente, apelando al medio ambiente, a menos que supongamos que cambios en las expectativas no tienen causas.

Esta consideración se hace patente en la siguiente situación, planteada por Guthrie (1933). Supóngase que se nos pide *predecir* cómo un organismo se comportará ante ciertas condiciones. *¿Qué información exigiríamos como base para tal predicción?* Cualquier psicólogo (incluyendo los psicólogos cognoscitivos) solicitaría, como mínimo, información sobre hechos que pueden ser establecidos mediante observaciones pasadas y/o presentes del comportamiento y el *ambiente* del organismo en cuestión. Sin embargo, de nuevo, por razones que desconozco, los psicólogos cognoscitivos tienden a considerar al ambiente como un mero mediador que juega un papel causal menor, relativo a aquel jugado por los estados y procesos cognoscitivos. En las explicaciones que he construido, el ambiente juega un papel causal tan central como el que puedan jugar tales procesos.

De cualquier modo, no cabe la menor duda de que analizar *O* cognoscitivamente resulta en propuestas más elaboradas que aquellas que construí antes de dicho análisis. Sin embargo, ¿constituyen tales propuestas explicaciones epistémicamente legítimas bajo los modelos ND y RE? Una respuesta instintiva a esta pregunta pareciera ser obvia y directa, cuando nos preguntamos por el significado del predicado *E*. En efecto, tal significado no puede ser conductual, al menos en referencia al *mismo* nivel de análisis en el cual se utilizan los predicados *P* y *R*. De otro modo, *E* sería sinónimo de *P* y *R*, por lo cual los enunciados resultantes serían analíticamente verdaderos (i.e., verdaderos *exclusivamente* en virtud del significado de esos predicados). El predicado *E* podría referirse a niveles de análisis conductual que tienen lugar por debajo del nivel en el cual se utilizan *P* y *R*, en cuyo caso los enunciados resultantes serían sintéticos (i.e., verdaderos *parcialmente* en virtud del significado de sus términos). Sin embargo, el psicólogo cognoscitivo típico renuncia a tal interpretación, así como también a una interpretación *biológica*, en particular *neurocientífica*, con lo cual pareciera retirarle el carácter de ley *empírica* a los enunciados que formula. En consecuencia, sus propuestas parecieran no calificar como explicaciones legítimas bajo los modelos ND y RE, en tanto en cuanto éstos *parecieran* exigir que las leyes que constituyen el *explicans* de una explicación genuina sean empíricas.

En efecto, algo pareciera calificar como una explicación científica epistémicamente legítima bajo ambos modelos si las leyes y condiciones iniciales que constituyen el *explicans* son *justificadas independientemente* unas de otras de manera *empírica*. Popper (1972), por ejemplo, expresa este requisito claramente:

“Para ser satisfactorio (y ello puede ser una cuestión de grado), el *explicans* debe satisfacer ciertas condiciones. Primero, debe entrañar lógicamente al *explicandum*. Segundo, el *explicans* debe ser verdadero; en cualquier caso, no debe ser falso, aun luego del examen más crítico. Si no se ha establecido como verdadero (lo que usualmente será el caso), debe haber evidencia *independiente* a su favor. En otras palabras, debe ser *independientemente* contrastable; y lo consideraremos como más satisfactorio mientras mayor sea la severidad de las pruebas que ha sobrevivido” (Popper, p. 192).

En la segunda condición, Popper expresa lo que podríamos denominar ‘*criterio de contrastabilidad independiente*’ de las leyes empíricas. Según este criterio, una condición mínima necesaria (mas no suficiente, por supuesto) para que un enunciado califique como una ley empírica es que debe permitir la formulación de condiciones que permitan una *contrastación empírica independiente*. Igualmente, Hempel y Oppenheim (1948), por su parte, plantean que “el explanans debe tener contenido empírico: i.e., debe ser capaz, al menos en principio, de prueba mediante la experimentación u observación” y que “los enunciados que constituyen el explanans deben satisfacer alguna condición de corrección fáctica” (p. 137).

Cualquier enunciado que esté constituido por *O*, *P*, y *R* (e.g.,  $(\forall x)[(Ox \wedge Px) \rightarrow Rx]$ ) ciertamente satisface este criterio, ya que tales predicados se refieren a propiedades conductuales observables más o menos de manera directa. Pero, ¿qué decir de los enunciados en los que participa *E* (o cualquier otro predicado cognoscitivo, si a ver vamos)? ¿Califican tales enunciados como leyes ‘empíricas’? Una respuesta negativa pareciera inevitable, puesto que *E* no es interpretado ni conductual ni neurobiológicamente, lo cual impide una contrastación empírica directa e independiente. De hecho, el psicólogo cognoscitivo típico aceptaría que aquellos procesos y eventos a los cuales se refiere *E* son *en principio inobservables de manera directa*. Lo único que podemos hacer, reconocería, es *inferir* tales procesos y eventos indirectamente a partir de relaciones entre la actividad observable de los organismos y su medio ambiente.

En el caso de la aplicación del modelo ND, entonces,  $(\forall x)[(Ox \wedge Px) \rightarrow Ex]$  y  $(\forall x)[(Ox \wedge Ex) \rightarrow Rx]$ , no expresarían leyes empíricas, al menos en el sentido de que puedan ser contrastadas independientemente mediante *observación directa*. Lo mismo puede decirse del conjunto de sentencias que resultó de aplicar el modelo RE. En efecto, las particiones de

$$O_1 = \{x | Ox \wedge Px\}$$

en

$O_{1,1} = \{x|Ox \wedge Px \wedge Ex\}$  y  $O_{1,2} = \{x|Ox \wedge Px \wedge \neg Ex\}$ ,  
y de

$$O_{1,1} = \{x|Ox \wedge Px \wedge Ex\}$$

en

$$O_{1,1,1} = \{x|Ox \wedge Px \wedge Ex \wedge Rx\}$$
 y  $O_{1,1,2} = \{x|Ox \wedge Px \wedge Ex \wedge \neg Rx\}$ ,

no pueden ser justificadas en términos de predicados que se refieren a propiedades conductuales directamente observables. Igualmente, la única forma de asignar valores de probabilidad a  $(Ox|Px \wedge Ex)$ ,  $\pi(Ox|Ox \wedge Px \wedge \neg Ex)$ ,  $\pi(Ox|Px \wedge Ex \wedge Rx)$ , y  $\pi(Ox|Ox \wedge Px \wedge Ex \wedge \neg Rx)$  sería postulando alguna distribución *teórica* de probabilidad que describa, por una parte, cambios en la probabilidad de **E** en función de **P**, y, por otra, cambios en la probabilidad de **R** en función de cambios de **E**.

#### *Los predicados cognoscitivos como términos teóricos*

Pareciera, pues, que según los modelos ND y RE, lo que los psicólogos cognoscitivos proponen como ‘explicaciones’ de la conducta, en realidad no representan explicaciones epistémicamente legítimas, ya que tales modelos aparentemente exigen que las leyes que se utilicen en el *explanans* sean empíricas y los enunciados que incluyen predicados cognoscitivos no califican como tales. Esta conclusión sería aceptable si y sólo si tales modelos admitieran como únicas leyes legítimas las leyes empíricas. Sin embargo, la cuestión se complica si tomamos en cuenta el análisis que Hempel y Oppenheim (1948) hacen acerca de las explicaciones en ciencias sociales y la propuesta de Salmon (1998) respecto a la causalidad.

En el primer caso, los autores arguyen que si bien su “caracterización de la explicación científica hasta ahora se basa en un estudio de casos tomados de las ciencias físicas”, dicha caracterización “también se aplica fuera de esta área”. “De esta manera”, continúan los autores,

“diversos tipos de conducta en animales de laboratorio y en sujetos humanos son explicados en psicología por la subsunción bajo leyes o aun teorías de aprendizaje o condicionamiento; y aun cuando frecuentemente las regularidades invocadas no pueden ser expresadas con la misma generalidad y precisión que en física o química, resulta al menos claro que el carácter general de tales explicaciones se adecua a nuestra caracterización anterior” (p. 140).

Aquí, los autores admiten la posibilidad de que una *teoría* puede tener poder explicativo (i.e., formar parte del *explicans* de una explicación). Ello se hace aun mas patente cuando los autores intentan extender su caracterización de la explicación al ámbito de la conducta humana. En efecto, luego de analizar brevemente ejemplos de explicaciones provenientes de la economía, la sociología y la lingüística, los autores se concentran en explicaciones psicológicas de la conducta intencional humana que apelan a ‘deseos’, ‘creencias’, ‘motivos’ y ‘propósitos’, arguyendo que:

“...el hecho de que los motivos no sean accesibles a observación directa por parte de un observador externo [no] constituye una diferencia esencial entre [explicaciones motivacionales en psicología, por una parte, y explicaciones causales en física y química, por otra]; puesto que también los factores determinantes aducidos en las explicaciones físicas son muy frecuentemente inaccesibles para la observación directa. Este es el caso, por ejemplo, cuando cargas eléctricas positivas son aducidas en la explicación de la atracción mutua de dos esferas metálicas. La presencia de tales cargas, aun cuando eluden la observación directa, pueden ser determinadas mediante diversos tipos de pruebas indirectas, y esto es suficiente para garantizar el carácter empírico del enunciado explicativo. De manera similar, la presencia de ciertas motivaciones puede ser determinada sólo mediante métodos indirectos, los cuales pueden incluir referencias a expresiones lingüísticas por parte del sujeto en cuestión, deslices de escritura o de la lengua, etc.; pero siempre que estos métodos estén ‘operacionalmente determinados’ con claridad y precisión razonables, no hay diferencia esencial entre explicación motivacional y explicación causal en física.” (p. 143).

Los autores mismos del modelo ND, entonces, aceptan como epistémicamente legítimas explicaciones que incluyen predicados cognoscitivos. Y al otorgar un poder explicativo a enunciados que incluyen tales predicados, están considerándolos como objetivamente existentes, más que como meros instrumentos de predicción o descripción. Estos autores, entonces, aceptarían, los enunciados  $(\forall x)[(Ox \wedge Px) \rightarrow Ex]$  y  $(\forall x)[(Ox \wedge Ex) \rightarrow Rx]$  como epistémicamente genuinos en tanto en cuanto son susceptibles de *verificación empírica indirecta*. De hecho, los autores considerarían enunciados constituidos exclusivamente por predicados cognoscitivos (e.g., las ecuaciones del modelo RW, que no incluyen predicado conductual alguno) como igualmente genuinos, siempre y cuando puedan ser relacionados lógicamente con enunciados que expresen ciertas operaciones empíricas claras y precisas, no importa cuán indirecta sea esa relación.

Tal aceptación por parte de Hempel y Oppenheim representa una de las justificaciones más frecuentemente esgrimidas por los psicólogos cognoscitivos para introducir predicados cognoscitivos y construir con ellos enunciados legaliformes que,

se supone, poseen poder explicativo. El argumento típico al respecto, por parte de los psicólogos cognoscitivos, es que ellos tienen el mismo derecho que los físicos de postular entidades, eventos y procesos inobservables como recursos explicativos. Y si hasta ahora tal estrategia ha funcionado en la física, pues también debería funcionar en la psicología. Hempel y Oppenheim continúan elaborando esta idea, de la siguiente manera:

... con frecuencia se hace una distinción entre varios niveles de explicación. La subsunción de un fenómeno bajo una ley general que conecta directamente características observables representa el primer nivel; niveles más altos requieren el uso de constructos teóricos más o menos abstractos, los cuales funcionan en el contexto de alguna teoría comprensiva. ... el concepto de explicación de nivel superior cubre procedimientos de un carácter muy diferente; uno de los más importantes entre ellos consiste en explicar una clase de fenómenos mediante una teoría acerca de su micro-estructura. La teoría cinética del calor, la teoría atómica de la materia, las teorías tanto electromagnética como cuántica de la luz, y la teoría genética de la herencia son ejemplos de este método. Con frecuencia, se siente que sólo el descubrimiento de una micro-teoría permite un entendimiento científico real de cualquier tipo de fenómeno, porque sólo ella nos da un 'insight' en el mecanismo interno del fenómeno, por así decirlo. (p. 147).

En este párrafo, los autores expresan una de las ideas centrales del empirismo lógico, a saber, la dicotomía observacional-teórico<sup>4</sup>. Según esta dicotomía, el vocabulario de la ciencia se divide en dos tipos de términos mutuamente excluyentes, a saber, términos *teóricos* y términos *observacionales*. Un enunciado que está constituido exclusivamente por términos teóricos se denomina 'enunciado teórico', mientras que uno constituido sólo por términos observacionales se denomina 'enunciado observacional'. Las leyes empíricas (entendidas como *generalización inductivas* a partir de ciertas *regularidades* observadas de manera más o menos directa) que hemos considerado hasta ahora califican, bajo esta dicotomía, como enunciados observacionales. Hasta ahora no he dado ejemplo alguno de una ley teórica, pero, en el caso que nos ocupa, las ecuaciones del modelo RW calificarían como leyes teóricas.

<sup>4</sup> Esta dicotomía (así como la concepción de las teorías científicas que de ella se deriva), tal y como fue propuesta por los positivistas lógicos, adolece de serios problemas (e.g., Moulines, 1982; Putnam, 1962; Stegmüller, 1973; Suppe, 1977; 1989). El hecho de que la esté adoptando aquí, entonces, no necesariamente significa que la acepto como correcta. Simplemente, estoy tratando de ser consistente con el modelo ND, el cual presupone dicha dicotomía. De hecho, Salmon también la presupone en su aproximación a la causalidad (ver más adelante), a pesar de las graves dificultades lógicas que, como se ha demostrado ampliamente, presenta. Entonces, la dicotomía lógico-positivista entre lo observacional y lo teórico no ha sido superada en la filosofía, mucho menos en las ciencias empíricas (incluyendo a la psicología, tanto cognoscitiva como conductual). El presente análisis, entonces, podría ser fácilmente rechazado bajo una concepción alternativa de las teorías científicas (e.g., la concepción estructuralista, o la concepción semántica). Sin embargo, tal rechazo llevaría a un tipo de análisis muy diferente del que estoy desarrollando aquí.



Carnap (1966) planteó la dicotomía observacional-teórico de la siguiente manera:

Una de las distinciones más importantes entre dos tipos de leyes en ciencia es la distinción entre lo que pueden ser llamadas (no hay una terminología generalmente aceptada para ellas) leyes empíricas y leyes teóricas. Las leyes empíricas son leyes que pueden ser directamente confirmadas por observaciones empíricas. El término 'observable' es con frecuencia utilizado para cualquier fenómeno que puede ser directamente observado, así que se puede decir que las leyes empíricas son leyes acerca de observables. .... Filósofos y científicos usan los términos 'observable' e 'inobservable' de muy diferentes maneras. Para un filósofo, 'observable' tiene un significado muy estrecho. Se aplica a propiedades tales como 'azul', 'duro', 'caliente'. Estas son propiedades directamente percibidas por los sentidos. Para el físico, la palabra tiene un significado mucho más amplio. Incluye cualquier magnitud cuantitativa que pueda ser medida de una manera relativamente simple, directa. Un filósofo no consideraría una temperatura de, quizás, 80 grados centígrados, o un peso de 98.5 libras, un observable porque no hay una percepción sensorial directa de tales magnitudes. Para un físico, ambas son observables porque pueden ser medidas de una manera extremadamente simple. (p. 147).

Más adelante, en el mismo libro, Carnap plantea que:

Si una magnitud permanece igual dentro de distancias espaciales o intervalos de tiempo lo suficientemente grandes, de tal manera que un aparato pueda ser aplicado para [obtener] una medición directa de la magnitud, se le denomina *macroevento*. Si la magnitud cambia dentro de intervalos de espacio y tiempo tan extremadamente pequeños que no puede ser medida a través de un aparato simple, es un *microevento*. ...Un microproceso es simplemente un proceso que involucra intervalos extremadamente pequeños de espacio y tiempo. Por ejemplo, la oscilación de una onda electromagnética de luz visible es un microproceso. Ningún instrumento puede medir directamente cómo esta intensidad varía. La distinción entre macro- y microconceptos es algunas veces tomada como un paralela a lo observable y lo inobservable. No es exactamente lo mismo, pero sí gruesamente. Las leyes teóricas conciernen no observables, y muy frecuentemente éstos son microprocesos. Si es así, las leyes son algunas veces llamadas 'microleyes'. Yo uso el término 'leyes teóricas' en un sentido más amplio que este, para incluir todas aquellas leyes que contienen no observables, independientemente de que sean microconceptos o macroconceptos. (pp. 227-228).

Según la formulación de Carnap, entonces, ciertos términos teóricos (o 'inobservables') involucran microconceptos. Por consiguiente, ciertas leyes teóricas

son microleyes. La razón principal por la cual los procesos y entidades a las cuales aluden tales términos y leyes se consideran como 'inobservables' es que, precisamente, tienen lugar en escalas espacio-temporales demasiado pequeñas como para ser directamente accesibles a través de la percepción humana o para ser medidos directamente a través de aparatos sencillos de utilizar. Esta posición supone una *ontología jerarquista* según la cual los constituyentes básicos de la realidad (sean universales, sean particulares, o ambos) pueden ser clasificados en *niveles de organización*, algunos de los cuales son perceptualmente accesibles para los humanos, mientras que otros no lo son<sup>5</sup>.

Bajo esta posición, para explicar ciertas leyes empíricas (y, en esa medida, *ciertas* ocurrencias particulares) resulta *inevitable* postular la existencia *microentidades* (e.g., partículas subatómicas) y la ocurrencia de *microprocesos* (e.g., oscilaciones de campos electromagnéticos) que no pueden ser directamente observados, ni siquiera en principio, pero que, a pesar de ello, pueden adquirir un contenido empírico en la medida en que puedan ser ligados a operaciones de medida clara y precisamente definidas, no importa cuán indirectas y complejas sean. Por supuesto, la forma lógica de una explicación ND de una ley empírica sería muy diferente de la presentada arriba, por cuanto su *explicandum* sería el enunciado  $(\forall x)[(Ox \wedge Px) \rightarrow Rx]$  (en lugar del enunciado  $Ra$ ), mientras que el *explicans* incluiría los enunciados  $(\forall x)[(Ox \wedge Px) \rightarrow Ex]$  y  $(\forall x)[(Ox \wedge Ex) \rightarrow Rx]$ . De hecho, estrictamente hablando, estos dos últimos enunciados serían considerados bajo esta concepción más bien como componentes de una *regla de correspondencia*, en lugar de una ley teórica, puesto que consisten tanto de un término teórico como de términos observacionales.

El papel principal de las reglas de correspondencia en esta concepción era de servir de 'puente' o 'enlace lógico' entre las leyes teóricas (constituidas sólo por predicados teóricos) y leyes empíricas (constituidas sólo por predicados observacionales). Aplicando la forma lógica apropiada para tales reglas, acordada por los empiristas lógicos, resultaría en el enunciado  $(\forall x)[(Ox \wedge Px) \rightarrow (Ex \leftrightarrow Rx)]$ <sup>6</sup>. Además de estos enunciados, por supuesto, una explicación ND de aquella ley empírica debería incluir leyes teóricas. Tal y como lo mencioné anteriormente, en el caso que nos ocupa, tales leyes vendrían dadas por las ecuaciones del modelo RW, las cuales (adecuadamente expresadas en el lenguaje de la lógica de predicados), en conjunción con aquella regla

<sup>5</sup> Según Carnap, un objeto, evento o proceso puede también ser inobservable debido a que ocurre en escalas espacio-temporales demasiado *grandes*. Pero aquí me concentraré sobre micro-conceptos.

<sup>6</sup> La adopción de esta forma para las reglas de correspondencia obedece a razones lógicas que no viene al caso elaborar aquí. Cabe señalar, sin embargo, que el modelo RW no incluye regla de respuesta alguna, es decir, una regla que permita transformar cambios cuantitativos en la fuerza asociativa en cambios cuantitativos en la respuesta. Bajo la concepción lógico-positivista, una teoría es una conjunción de un conjunto de leyes teóricas y un conjunto de reglas de correspondencia. Estrictamente hablando, entonces, el modelo RW, visto bajo dicha concepción, es una *pseudoteoría*, ya que no incluye regla de correspondencia alguna que sea clara y precisa. En ausencia de una regla de correspondencia, esta pseudoteoría puede ser evaluada sólo mediante juicios comparativos extremadamente gruesos, intuitivos y visuales.

de correspondencia, permitirían, en principio, la deducción de la ley empírica en cuestión. El argumento deductivo resultante vendría a constituir una explicación *bona fide* ND de la ley empírica en cuestión. Sobre esta base, además, podríamos explicar el hecho expresado por el enunciado *Ra* en términos de leyes teóricas, adjuntando la explicación de la ley empírica a la explicación de ese hecho, de tal manera que éste sea la conclusión y la ley empírica constituya un paso deductivo derivado de las leyes teóricas y de la regla de correspondencia.

Por su parte, Salmon (1998) resume su posición actual respecto a su propio modelo de la siguiente manera:

Si deseamos explicar un evento particular, ..., empezamos ensamblando los factores que son estadísticamente relevantes para la ocurrencia. ... Debemos también obtener los valores de probabilidad asociados con las relaciones de relevancia. Las relaciones de relevancia estadística son regularidades estadísticas, y procedemos a explicarlas. ..., ya no creo que el ensamblaje de los factores relevantes provea una explicación completa. ... Creo que tenemos una explicación *bona fide* de un evento si tenemos un conjunto completo de factores estadísticamente relevantes, los valores de probabilidad pertinentes, y explicaciones causales de las relaciones de relevancia. Subsunción de una ocurrencia particular bajo regularidades estadísticas -la cual, como hemos visto, no implica nada acerca de la construcción de argumentos deductivos o inductivos- es una parte necesaria de cualquier explicación adecuada de su ocurrencia, pero esa no es toda la historia. La explicación causal de la regularidad también es necesaria. Esta aserción, debe notarse, está en conflicto directo con la concepción heredada, de acuerdo con la cual la mera subsunción -deductiva o inductiva- de un evento bajo una regularidad legal constituye una explicación completa. Uno puede, de acuerdo con la concepción heredada, pedir una explicación de cualquier ley usada para explicar un evento dado, pero eso es una explicación diferente. Estoy sugiriendo, por el contrario, que si la regularidad invocada no es una regularidad causal, entonces la explicación causal de esa regularidad debe hacerse parte de la explicación del evento (p. 137).

En este párrafo, por 'concepción heredada', Salmon se refiere, por supuesto, al modelo ND (respecto a la relación de subsunción deductiva) y a la versión estadístico-inductiva del mismo (respecto a la subsunción inductiva), la cual fue posteriormente propuesta por Hempel (1965) para caracterizar explicaciones basadas en leyes probabilísticas. Este cambio de pensamiento expresado por Salmon respecto a su propio modelo obliga a retirar el predicado *E* de la segunda explicación RE que presenté arriba, para regresar a la primera explicación, basada sólo en los predicados *O*, *P*, y *R* por cuanto, de nuevo, sólo éstos permiten la formulación de regularidades estadísticas

estrictamente empíricas. Lo que Salmon propone, entonces, es que esa explicación original debe ser *complementada* por un conjunto de *leyes causales* que expliquen las regularidades empíricamente determinadas como estadísticamente relevantes.

En su concepto de lo que es una ley causal, Salmon (1984, 1998) adopta completa e incondicionalmente la dicotomía observacional-teórico:

Aunque podemos dotar al mundo con mucha continuidad mediante la referencia a entidades observables (aunque inobservadas), no podemos hacer un trabajo muy completo al respecto. Para realizar la tarea, debemos introducir algunas entidades que no son observables, al menos por las capacidades humanas ordinarias de percepción. Si, por ejemplo, notamos que las ventanas de la cocina tienden a ahumarse en los días fríos cuando estamos hirviendo agua en la estufa, conectamos la hervida del agua en la estufa con la ahumada de las ventanas hipotetizando la existencia de moléculas de agua que son demasiado pequeñas para poder ser vistas por el ojo desnudo, y afirmando que viajan en trayectorias continuas desde la olla hasta la ventana. Consideraciones similares llevaron a la postulación de microbios, virus, y genes para la explicación de fenómenos tales como el esparcimiento de una enfermedad y la herencia de características biológicas. ... Las ondas electromagnéticas son invocadas para cumplir el mismo tipo de función; en la explicación de la perturbación de la señal de TV, la propagación de ondas electromagnéticas provee la conexión continua. (pp. 227-228).

Para Salmon, entonces, una ley causal es un enunciado teórico en el cual se postula una conexión real y necesaria entre micro-procesos y macro-procesos, donde los primeros subyacen a los segundos. De este modo, la noción de ley causal de Salmon se cuadra completamente con la dicotomía tradicional entre lo teórico y lo observacional.

En suma, la estrategia cognoscitiva de postular procesos y estados inobservables para explicar la conducta observable de los organismos pareciera ser perfectamente legítima tanto desde la perspectiva del modelo ND como de la del modelo RE. Ante tal legitimación filosófica, ¿cómo queda el rechazo del conductismo radical hacia dicha estrategia? Por supuesto, si a ver vamos, a los conductistas radicales típicos les tiene sin cuidado lo que piense el filósofo de la ciencia respecto a la psicología cognoscitiva o al conductismo radical. Ante tal legitimación, el grueso de los conductistas radicales tenderá a reaccionar concluyendo que el filósofo de la ciencia está tan equivocado como el psicólogo cognoscitivo y, por tanto, a verlos a ambos como enemigos intelectuales mercedores de la crítica y el rechazo más severos. Después de todo, continuaría el conductista radical típico (no sin cierta razón), los filósofos no tienen la última palabra respecto a nada, por lo cual podemos perfectamente ignorar lo que nos dicen y continuar tranquilamente con nuestro trabajo. Sin embargo, para bien o para mal, las cosas no

son tan sencillas. Antes de examinar la situación más cuidadosamente, veamos qué nos dice un tercer modelo de la explicación.

### *Inferencia a la mejor explicación*

Mediante el modelo de la IME se intenta resolver el problema de la inducción, describiendo nuestras prácticas de inferencia inductiva. Bajo este modelo, “nuestras prácticas explicativas guían nuestras inferencias” (Lipton, 1991, p. 1). El procedimiento básico consiste en inferir, a partir de cierta evidencia disponible, aquello que, *de ser verdadero*, nos permitiría *entender mejor* esa evidencia. Nótese que la condición de verdad en este modelo es *disposicional* o *potencial*, por lo cual una explicación no necesariamente tiene que ser verdadera para ser poderosa. Una explicación falsa puede muy bien ser mejor que una explicación verdadera, si la primera nos permite ‘entender mejor’ la evidencia disponible.

A manera de ejemplos, en presencia de huellas con forma de raqueta en la nieve, tendemos a inferir que una persona con cierto tipo de calzado (raquetas de nieve) caminó por allí recientemente. Hay otras posibilidades, pero llevamos a cabo tal inferencia porque provee la mejor explicación. Cuando vemos a alguien retirar la mano repentinamente de una llama, inferimos que la persona ha experimentado dolor, porque esta es la mejor explicación de la conducta observada. El detective infiere que el mayordomo lo hizo, puesto que esta es la mejor explicación de la evidencia disponible. Inferimos que la noche anterior (mientras dormíamos profundamente) llovió, ya que es la mejor explicación de la presencia de agua en el suelo y en los árboles. El médico infiere que su paciente tiene paperas, ya que esta es la ‘mejor’ explicación de los síntomas. El astrónomo infiere la existencia y movimiento de Neptuno ya que es la ‘mejor’ explicación de las perturbaciones observadas en la órbita de Urano. Y así sucesivamente. En todos estos casos, las explicaciones respectivas pueden muy bien ser falsas. Pero eso no les quita su poder explicativo.

En el fenómeno que nos ocupa, este modelo nos permite inferir, en presencia de un incremento en el responder del organismo ante el EC luego de un tratamiento estándar de condicionamiento pavloviano, que la expectativa del organismo de que el EI ocurrirá luego del EC fue incrementada. De hecho, la práctica de inferir estados y procesos cognoscitivos (inobservables) a partir de la conducta observable de los organismos es también vista como perfectamente legítima, natural y hasta deseable bajo el modelo IME, tal y como el mismo Lipton (1991) lo expresa:

“...con frecuencia hacemos inferencias inductivas de algo invisible a partir algo que observamos, como por ejemplo de las creencias de las personas a partir de su conducta o de entidades y procesos inobservables a partir de evidencia científica.” (p. 16).

“El método de la desviación explicativa parece ser una de las fuentes del gran poder predictivo y manipulativo de muchas áreas de la ciencia. En ciencia, la desviación con frecuencia exige inferencia ‘vertical’ a explicaciones en términos de entidades y procesos inobservados y con frecuencia inobservables, y la Inferencia a la Mejor Explicación parece bien equipada para dar cuenta de este proceso.” (p. 68).

Al igual que la versión revisada del modelo RE por parte de Salmon, el modelo de la IME (tal y como es presentado por Lipton) también supone una relación estrecha entre explicación y causalidad:

“... explicar un fenómeno es simplemente proveer información acerca de su historia causal ... o, si el fenómeno mismo es una regularidad causal, explicar es dar información acerca del mecanismo que relaciona causa con efecto.” (p. 32).

“... propongo que, para las explicaciones causales de eventos, los contrastes explicativos seleccionan causas por medio de lo que denominaré la ‘Condición de Diferencia’. *Para explicar por qué P en lugar de Q, debemos citar una diferencia causal entre P y no-Q, que consista de una causa de P y la ausencia de un evento correspondiente en caso de que no-Q.*” (p. 43).

Si combinamos estas dos citas con las dos anteriores, parece inevitable concluir (al igual que con los modelos ND y RE) que al menos ciertas causas son inobservables. Bajo este modelo, entonces, observabilidad directa no es un requisito para que algo califique como causa. Ello es consistente con el hecho de que, bajo este modelo, la inferencia a *cualquier* explicación adquiere sentido sólo en aquellos casos en los cuales carecemos de cierta información independiente y directa, de tal manera que nuestra única opción es elegir la mejor explicación, *de un conjunto de explicaciones posibles*. Por ejemplo, si hubiéramos visto a alguien con raquetas de nieve caminando por la nieve, pues obviamente no necesitamos inferir nada ante las huellas, puesto que ya poseemos la información crítica que las explica. Asimismo, si el mayordomo confiesa, pues no necesitamos considerar un conjunto de posibles explicaciones y elegir la ‘mejor’. Si el planeta Neptuno hubiera sido descubierto mediante observación telescópica directa, pues no hubiera sido necesario inferir su existencia. Si hubiéramos estado despiertos la noche anterior y hubiéramos visto y escuchado llover, no hubiera sido necesario inferir

que la noche anterior llovió. Lo mismo se aplica a la inferencia de causas que son en principio inobservables, debido a que ocurren en escalas espacio-temporales demasiado pequeñas o demasiado grandes.

Lo que según el modelo de la IME nos obliga a hacer inferencias inductivas (guiados por consideraciones explicativas), pues, es la *carencia* de cierta información crítica que relacione de alguna forma coherente y significativa (i.e., explicativa) las distintas piezas de cierta evidencia disponible, de tal manera que logremos un entendimiento de la situación. Inferir la mejor explicación, pues, tiene su origen en nuestra *ignorancia* y es, por tanto, epistémicamente relativo a un cuerpo de evidencia disponible. Técnicamente, nuestra situación de ignorancia se expresa diciendo que la evidencia disponible (más las reglas de deducción lógica) *subdetermina* nuestras inferencias (i.e., no garantiza una solución o respuesta única al problema planteado).

En el tipo de fenómeno que he estado considerando, son posibles al menos dos formas no triviales de subdeterminación. Primero, si poseemos evidencia solamente del responder del organismo bajo ciertas condiciones actualmente observadas, entonces enfrentamos el problema de inferir la historia. Es decir, si la única evidencia disponible es que un organismo saliva ante un sonido, entonces inferiríamos que el organismo fue sometido a un tratamiento estándar de condicionamiento pavloviano, sobre la base de que ésta sería la mejor explicación de tal evidencia. Segundo, si sólo poseemos evidencia de la historia y de la conducta actual, entonces se nos plantearía el problema de inferir qué sucede dentro del organismo durante dicha historia. En ambas situaciones, cierta evidencia disponible subdetermina (i.e., determina de manera multivoca) soluciones a ciertos problemas o preguntas que nos planteamos. También, en ambos casos, las inferencias que nuestros intentos de solución nos obligan a hacer son de carácter inductivo, por cuanto involucran inferencias de lo observado (la evidencia disponible) a lo no observado. Se trata, sin embargo, de inferencias inductivas *heterogéneas* o *liberales*, en el sentido de que lo observado y lo inferido involucran estructuras, procesos y eventos de diferentes tipos.

La estrategia de inferir estados y procesos cognoscitivos se basa sobre la suposición de que la información que poseemos en el segundo tipo de situación subdetermina nuestras hipótesis acerca de lo que ocurre dentro del organismo cuando su conducta cambia en función de cambios en el medio ambiente. Este es el único tipo de subdeterminación en la cual tendría sentido inferir procesos y eventos cognoscitivos. El psicólogo cognoscitivo típico, entonces, argüirá que la evidencia disponible es insuficiente para solucionar de manera unívoca el problema de qué sucede dentro del organismo durante el tratamiento. Se plantea, entonces, una situación en la cual nos vemos obligados a *inferir* una solución a dicho problema, asumiendo, claro está, la legitimidad del mismo. Esta forma de plantear el problema se basa sobre una distinción

entre afuera y adentro, externo e interno, la cual puede ser muy discutible. En este contexto, sin embargo, tomo tal distinción meramente como una distinción estrictamente *epistemológica* entre *niveles de análisis* que involucran conceptos diferentes y requieren de procedimientos diferentes para ser estudiados, sin connotación metafísica alguna. Con ello no estoy meramente concediéndole al psicólogo cognoscitivo la legitimidad de tal distinción. Por el contrario, la considero como perfectamente legítima. De hecho, considero como perfectamente legítimo preguntar qué sucede dentro del organismo cuando éste se comporta. Ello no necesariamente significa que acepte la estrategia del psicólogo cognoscitivo, ni siquiera la posibilidad de que esta pregunta pueda ser adecuadamente respondida con los métodos de la psicología (aunque sean conductuales).

Al igual que los modelos ND y RE, el modelo de la IME pareciera legitimar la estrategia cognoscitiva de explicar la conducta observable en términos de procesos y estados internos inferidos a partir de dicha conducta. Pareciera, entonces, que desde el punto de vista de tres modelos de la explicación, el psicólogo cognoscitivo tiene todo el respaldo y el conductista radical ninguno. Pareciera que los conductistas radicales se encuentran solos en este campo de batalla. De nuevo, los conductistas radicales típicos podrían reaccionar ante esta situación adoptando una actitud de desdén hacia el resto del mundo, alimentándose unos a otros la convicción de que son ellos los que tienen la razón, asumiendo en ocasiones la posición de mártires que habitan un mundo académico irracional y absurdo que se niega a ver la luz. Pero, de nuevo, las cosas no son tan sencillas. A continuación, examino el problema más cuidadosamente y propongo un argumento en contra de la estrategia cognoscitiva más elaborado del que usualmente ofrecen los conductistas radicales.

## UN ARGUMENTO MÁS ELABORADO

### *La posición skinneriana acerca de las explicaciones cognoscitivas*

Empecemos preguntándonos sobre qué bases el conductista radical típico rechazaría la estrategia cognoscitiva. Un argumento típico se basa sobre la inobservabilidad (o inaccesibilidad empírica) en principio de tales procesos o estados (e.g., Chiesa, 1994). El argumento básico en cuestión es que la accesibilidad empírica es la piedra angular de toda ciencia experimental. Lo que no es empíricamente accesible, entonces, cae fuera del ámbito de la experimentación y, por tanto, no es susceptible de ser estudiado (mucho menos entendido) científicamente. Este argumento se extiende a cualquier intento de explicar lo empíricamente accesible (la conducta observable) en términos de lo empíricamente inaccesible (estados y procesos inferidos a partir de la conducta observable). De nuevo, según este argumento, si lo que se propone como *explicans*



involucra inobservables, entonces no es posible determinar si éste es empíricamente verdadero o falso. Sólo lo empíricamente verdadero posee poder explicativo. Por lo tanto, los enunciados teóricos no poseen valor explicativo alguno.

Toda la discusión anterior debería dar suficientes razones para concluir que tal argumento resulta, por demás, ingenuo, superficial y flojo. En efecto, resulta obvio que el uso de inobservables como recursos explicativos (lo cual implica, de nuevo, postular su existencia objetiva) representa una práctica lo suficientemente extendida en las diversas ramas de la física (especialmente en el electromagnetismo, la mecánica cuántica y la termodinámica) como para rechazarla tan ligeramente como los conductistas radicales tienden a hacerlo<sup>7</sup>. Al rechazar la estrategia cognoscitiva, el conductista radical tendría igualmente que rechazar la estrategia de los físicos, a menos que elabore su argumentación de tal manera que le permita hacer lo primero sin necesidad de hacer lo segundo (esto es lo que precisamente intentaré aquí; ver más adelante).

Su otra posibilidad, por supuesto, sería rechazar con igual vehemencia la estrategia de los físicos (ver Nota 7). Sin embargo, tal rechazo sería lo suficientemente drástico como para requerir una justificación mucho más convincente (clara, precisa, sistemática e internamente coherente) de las que cualquier conductista radical ha ofrecido hasta ahora. Además, el mismo Skinner (1974) explícitamente rechazó el argumento de la inobservabilidad, al separar la forma de conductismo que él defendió (el conductismo "radical") del conductismo "metodológico". La idea básica de Skinner al respecto es que no debemos negar la existencia de estados y procesos internos por el mero hecho de que son inobservables mediante los procedimientos de la psicología conductual (como haría el conductismo metodológico). Skinner (1950, 1977, 1989) más bien rechazó la estrategia cognoscitiva sobre bases metafísicas y pragmáticas.

En lo metafísico, Skinner cuestionó, por una parte, la *naturaleza inmaterial* de estados y procesos 'mentales', considerándolos más bien como de naturaleza fisiológica (y, por tanto, ajenos al ámbito conceptual y metodológico de una ciencia de la conducta) y, por otra, el *papel causal* de dichos estados y procesos (aun viéndolos como fisiológicos), sobre la base de la evidencia (por demás contundente) de que el ambiente

<sup>7</sup> Una excepción fue Mach (1893/1960), quien, como sabemos, influyó a Skinner (1938). Mach era abiertamente instrumentalista, por cuanto consideraba que no debemos adjudicar un papel explicativo a los inobservables. Sin embargo, a pesar de que Mach indudablemente representó una influencia decisiva sobre el empirismo lógico, su posición instrumentalista fue considerada como un tanto extrema por la mayoría de los empiristas lógicos, quienes, como he mostrado, tendían a atribuir un poder explicativo a los enunciados teóricos y, en esa medida, a considerarlos como verdaderos. Por otra parte, debe reconocerse que la rama de la física conocida como 'mecánica cuántica' (en particular, la llamada 'Interpretación de Copenhagen') ha tendido a ser fuertemente instrumentalista. No obstante, algunos físicos (e.g., Bohm y Hiley, 1993) y filósofos (e.g., Maxwell, 1993) han argüido que una interpretación ontológica realista de la mecánica cuántica no sólo es necesaria para resolver ciertos problemas conceptuales crónicos que han aquejado a la teoría desde sus inicios, sino también para generar predicciones que una interpretación estrictamente epistemológica o instrumentalista no es capaz de generar.

externo parece jugar un papel causal importante en la conducta. Al respecto, Skinner consideró que, como el comportamiento, tales estados y procesos eran *efectos* del ambiente. En lo pragmático, Skinner cuestionó la necesidad de tales procesos y estados (de nuevo, conceptuados como de naturaleza fisiológica) para la descripción, predicción y cambio efectivo y eficiente de la conducta.

A primera vista, la crítica metafísica parece clara e indiscutible. Sin embargo, no resiste un análisis medianamente inquisitivo. Por una parte, ¿exactamente qué significa 'material'? ¿Qué está constituido por materia? Si aceptamos esta definición, entonces ¿qué es materia? ¿Lo que ocupa espacio, posee tamaño y forma, masa, movimiento y solidez? Si aceptamos esta definición de lo que es materia, entonces la conducta, como el objeto de estudio del análisis conductual, es tan 'inmaterial' como los estados y procesos cognoscitivos, en tanto en cuanto la conducta ciertamente carece de masa, tamaño y forma, elementos definitorios de al menos un concepto intuitivo de materia. La adopción de conceptos menos intuitivos y más acordes con la física actual, en la cual la materia misma se resuelve en fuerzas y energías, resulta aún más problemática, puesto que nos obligaría a especificar exactamente en qué sentido la conducta se resuelve en fuerzas y energías. Por otra parte, respecto al papel causal de los estados y procesos cognoscitivos, habría que empezar por aclarar qué es una causa. Por el mero hecho de no hacerlo, este aspecto de la crítica metafísica de Skinner se convierte en una cuestión virtualmente muda.

Por su parte, la crítica pragmática, siendo mucho más frecuente en Skinner, es aun más problemática que la metafísica, por cuanto asume que la predicción y la descripción son necesarias y suficientes en una ciencia de la conducta, suposición que se presenta como axiomática, sin justificación alguna más allá de sí misma, ante lo cual cabría preguntar ¿por qué son la predicción y la descripción suficientes en una ciencia de la conducta? Además, esta crítica se basa sobre un uso de los términos 'explicar' y 'describir' como sinónimos, el cual resulta por demás discutible, tal y como lo planteé en la introducción.

¿Será, entonces, que los conductistas están equivocados en sus críticas a la estrategia cognoscitivista? ¿Será que adoptar tal estrategia es, después de todo, la manera realmente científica de proceder? Mis respuestas a estas preguntas serán afirmativas en un sentido y negativas en otro. Afirmativas en el sentido de que las críticas de los conductistas a la estrategia cognoscitiva son demasiado flojas, ingenuas y superficiales, y de que la postulación de procesos y estados inobservables (o, al menos, inobservados en el momento de la explicación) constituye un aspecto crítico de explicar científicamente. Negativas en el sentido de que tales procesos y estados no tienen que ser necesariamente los que postula el psicólogo cognoscitivo, sino que pueden ser *neuronales*.

*Explicaciones neurocientíficas*

En efecto, reexaminemos lo que nos dice el modelo IME. De nuevo, bajo este modelo, la inferencia del tipo de explicaciones que encontramos en la psicología cognoscitiva se justifica sobre la base de que la solución al problema de qué sucede dentro de los organismos cuando se comportan está subdeterminada por nuestras observaciones de la conducta. De esto no hay duda alguna. Sin embargo, cabe preguntar si los estados y procesos cognoscitivos realmente representan la *mejor explicación* de la conducta observada. Una aplicación sensata (objetivamente comprensiva) de este modelo involucraría tomar en cuenta *todo el conjunto de posibles soluciones* (resultados, productos, desenlaces), *sin excepciones*, y escoger la mejor. Y entre tales posibles soluciones tenemos, sin lugar a dudas, aquellas que provienen de las neurociencias. Desde la perspectiva del modelo, restringir el conjunto de posibles soluciones sólo a aquellas que nos ofrece la psicología cognoscitiva resulta arbitrario, intelectualmente deshonesto e injustificable. Después de todo, las neurociencias también hablan acerca de lo que ocurre dentro de los organismos cuando se comportan bajo diversas condiciones ambientales.

El modelo IME aun carece de criterios claros y precisos para decidir si las explicaciones provenientes de las neurociencias son mejores o peores que las provenientes de la psicología cognoscitiva. Uno pensaría que las explicaciones provenientes de las neurociencias sobresaldrían como las más *plausibles* (i.e., las más sólidas y mejor justificadas, aunque no necesariamente las más simples y elegantes), al menos respecto al tipo de fenómeno que he examinado aquí. Sin embargo, Lipton (1991) hace una distinción entre explicaciones “plausibles” y explicaciones “adorables”, arguyendo que las primeras no necesariamente son las mejores. Según Lipton, la explicación más “plausible” es la que está mejor justificada, mientras que la explicación más “adorable” es aquella que “provee el mayor entendimiento” (p. 59). Lo ideal, por supuesto, es que la explicación que elijamos como la mejor sea la más plausible y la más adorable. Pero tales explicaciones son muy difíciles de obtener (Lipton ofrece la mecánica newtoniana como un ejemplo; ver p. 62), ante lo cual Lipton arguye que la explicación más adorable no necesariamente es la más plausible. Bajo este modelo, entonces, plausibilidad (o verdad) no es un criterio de poder explicativo.

Parece claro que bajo cualquier criterio que utilizemos, una explicación neurocientífica es científicamente más plausible que una explicación cognoscitiva. A pesar de esto, ciertos fenómenos podrían muy bien admitir explicaciones cognoscitivas como las más “adorables”. Ello puede deberse, principalmente, al hecho de que explicaciones neurocientíficas de ciertos fenómenos conductuales (sobre todo de la conducta compleja humana), pueden resultar demasiado engorrosas y técnicas y, en esa

medida, demasiado 'odiosas' como para siquiera ser tomadas en cuenta. Después de todo, 'entender' un fenómeno depende de cuán intuitiva y simple sea la manera en la cual lo expliquemos.

En efecto, por lo general, las explicaciones adorables son lo suficientemente simples e intuitivas como para permitirnos entender de manera eficaz y eficiente (sin muchas complicaciones y sin entrenamiento especial alguno) aquello que estamos explicando. Reconozco que explicaciones en términos de expectativas, creencias y deseos (aun de fenómenos tan simples como el que he considerado a lo largo del presente artículo) resultan mucho más intuitivas y simples que explicaciones neurocientíficas. Sin embargo, también hay que reconocer que una característica de las explicaciones *científicas* es que tienden a alejarse del sentido y la intuición comunes, lo cual las hace menos adorables para el lego, pero no necesariamente para el científico mismo. En efecto, la psicología aun posee mucho de sentido común, lo cual, aparte de representar un índice de inmadurez (si tomamos en cuenta que la tendencia en las ciencias ha sido, para bien o para mal, alejarse del sentido común), indudablemente determina cuáles explicaciones son adorables y cuáles son odiosas. Por su parte, una explicación neurocientífica puede no resultar tan simple e intuitiva como para ser suficientemente adorable y por tanto elegida como la 'mejor' explicación bajo el modelo IME. Sin embargo, ello no necesariamente significa que las explicaciones cognoscitivas sean *inherente, intrínseca o absoluta-mente* mejores que las explicaciones neurocientíficas.

Por su parte, los modelos ND y RE hacen mucho más énfasis que el modelo IME sobre la plausibilidad y justificación de las leyes teóricas utilizadas en nuestros *explicans*, por lo cual las explicaciones neurocientíficas deberían ser preferibles a las cognoscitivas. Respecto al modelo ND, entonces, propongo que inferir estados y procesos neuronales (en aquellos casos en los cuales tales estados y procesos no hayan sido observados) es más consistente con la concepción lógico-positivista de las teorías científicas que inferir estados y procesos cognoscitivos. Ello a pesar del análisis de Hempel y Oppenheim acerca de las explicaciones motivacionales. Al respecto, propongo que estos autores sucumbieron a la misma tentación a la cual sucumben la gran mayoría de los psicólogos, a saber, preferir una explicación intuitiva y simple, en lugar de una explicación estrictamente científica (más clara, detallada y precisa). Por supuesto, en la actualidad, las neurociencias se encuentran en un estadio sustancialmente más avanzado que durante la época en la cual Hempel y Oppenheim escribieron su artículo (aunque los conceptos centrales de neurona y sinapsis ya estaban perfectamente establecidos). Por ello, debemos reevaluar el análisis llevado a cabo por estos autores acerca de las explicaciones motivacionales, a la luz de los avances mostrados por las neurociencias desde finales de los años 40.

*Términos teóricos en las neurociencias*

Tal reevaluación se justifica sobre la base de dos consideraciones. En primer lugar, los términos que nombran los distintos aspectos de los estados y procesos neuronales (e.g., 'neurona', 'sinapsis', 'potencial de acción', 'ion', 'permeabilidad', 'neurotransmisor', 'receptor', 'transmisión sináptica', etc.) califican como teóricos, en el sentido de que nombran micro-conceptos que involucran el uso de aparatos relativamente complicados. En segundo lugar, las leyes que se refieren a tales estados y procesos neuronales son mucho más precisas, claras, y contrastables que las leyes que se refieren a estados y procesos cognoscitivos. En consecuencia, las primeras poseen un respaldo empírico considerablemente más sólido, directo y detallado que las segundas.

Lo mismo puede decirse del modelo RE, puesto que Salmon también acepta la concepción lógico-positivista de la dicotomía observacional-teórico. Bajo ambos modelos justificativos de la explicación científica, pues, las leyes que provienen de las neurociencias son, a todas luces, epistémicamente preferibles a aquellas que provienen de la psicología cognoscitiva. Además, la adopción del modelo RE sería inevitable en la medida que ciertas leyes neuronales son estocásticas (e.g., la teoría cuántica de la liberación de sustancias neurotransmisoras, teoría que supone que dicha liberación sigue una distribución binomial).

Los eventos y procesos neuronales, entonces, también califican como inobservables, aunque quizás en un sentido específico un tanto distinto del que encontramos en los inobservables de la física o, inclusive, de la psicología cognoscitiva misma. Para aclarar este punto, debemos antes que nada recordar que las neurociencias son un cúmulo de disciplinas experimentales y teóricas que han mostrado un grado progresivamente creciente de aislamiento (tanto en lo conceptual como en lo metodológico) debido a la especialización. Por ello, es importante elaborar exactamente a qué nos referimos cuando hablamos de 'inobservables' en las neurociencias.

Todos los conceptos centrales de la neurobiología *molecular* (e.g., receptor, neurotransmisor, ion) son micro-conceptos y, en esa medida, se refieren a objetos (y a los eventos y procesos que los involucran) que son inobservables en el mismo sentido fundamental que, según la concepción lógico-positivista de la dicotomía teórico-observacional, encontramos en los inobservables de ciertas áreas de la física, en particular la termodinámica, el electromagnetismo, y la mecánica cuántica. Esta es una forma común que, según esta concepción, adoptan los inobservables en ciencia, aunque, por supuesto, también hay macro-conceptos que involucran inobservables (e.g., un agujero negro). Pero en el caso de la neurociencia molecular, todos los inobservables son micro-conceptos, y la razón fundamental de que sean inobservables es que tienen lugar en escalas espacio-temporales demasiado pequeñas como para ser observados

directamente mediante los sentidos o mediante aparatos relativamente sencillos de usar.

Por su parte, el objeto de estudio de la neurobiología celular es la estructura y función de neuronas individuales. Tal estructura y función necesariamente incluye aquellos objetos y eventos propios de la neurobiología molecular, por lo cual ambas ramas de las neurociencias son difícilmente separables. Sin embargo, resulta claro que una neurona sólo puede ser observada mediante el uso de aparatos especiales (microscopios), los cuales pueden involucrar operaciones relativamente complicadas (especialmente en el caso de los microscopios electrónicos). Por consiguiente, las neuronas también califican como objetos inobservables, en el sentido de la concepción lógico-positivista de la dicotomía observacional-teórico.

Los conceptos de la neurobiología molecular y celular constituyen la base sobre la cual descansan las otras ramas de las neurociencias (neuroanatomía, neuroendocrinología, neuroinmunología, e inclusive la neurociencia conductual), y son enteramente micro-conceptos que se refieren a inobservables, en el sentido de que no pueden ser directamente observados a 'simple vista', sino indirectamente observados a 'compleja' o 'complicada' vista (i.e., mediante aparatos cuyo uso requiere de un grado considerable de entrenamiento).

Los conceptos de la psicología cognoscitiva son muy diferentes, por lo cual resulta difícil calificarlos de teóricos, al menos bajo la concepción lógico-positivista de los conceptos teóricos. En efecto, por una parte, es cierto que términos tales como 'expectativa', 'creencia', 'sorpresa', 'atención', 'fuerza asociativa', y otros de semejante guisa, supuestamente se 'refieren' a estados y procesos 'ocultos' en el sentido de que no pueden verse, escucharse, olerse, tocarse o saborearse directamente. Por otra parte, no obstante, la observación de sus supuestas 'manifestaciones descubiertas' tampoco requiere de operaciones muy sofisticadas que digamos. Las tres variables dependientes más utilizadas por los psicólogos cognoscitivos, a saber, autoreportes, tiempos de reacción (o latencias), y porcentajes de respuestas de cierto tipo, pueden ser estudiadas a través de operaciones que son considerablemente más sencillas y directas que aquellas que están involucradas en obtener el peso de un electrón, la intensidad de un campo electromagnético, la permeabilidad de una membrana, la concentración de iones de cierto tipo en el soma de una neurona, o un potencial de acción. Además, las relaciones entre tales variables y los estados y procesos que supuestamente les subyacen son rutinariamente ambiguas y oscuras. La situación es muy diferente en el caso de la física, la química, y la neurobiología molecular/celular, donde las relaciones entre los observables y los inobservables son considerablemente más precisas y claras.

Aunado a todo lo anterior, tenemos el problema de que no queda claro exactamente en qué sentido los estados y procesos cognoscitivos son inobservables. En el caso de la física, un objeto, estado o proceso es inobservable debido a que tiene lugar en escalas

espacio-temporales demasiado pequeñas (o demasiado grandes), mientras que en el caso de la neurobiología molecular/celular, un objeto estado o proceso es inobservable porque tiene lugar en escalas espacio-temporales demasiado pequeñas. De hecho, en ambos casos se declaran explícita, precisa y claramente (i.e., cuantitativamente) las escalas espacio-temporales en las cuales tienen lugar los distintos inobservables. Sin embargo, ningún psicólogo cognoscitivo hasta ahora ha especificado las escalas espacio-temporales en las cuales ocurren los estados y procesos que postula como explicaciones.

Pareciera, entonces, que tales estados y procesos son inobservables sólo en el sentido de que están ‘ocultos’ o ‘escondidos’. Pero si esto es así, ¿exactamente qué los oculta? Si quisiéramos encontrarlos, ¿dónde buscaríamos? Entonces, los estados y procesos cognoscitivos son ‘inobservables’ (y, en esa medida, los términos correspondientes son ‘teóricos’) en un sentido lo suficientemente diferente del que encontramos en la caracterización lógico-positivista de la dicotomía observacional/teórico, como para cuestionar el razonamiento de que si los físicos infieren tales estados y procesos, pues por qué no hemos de hacerlo también nosotros. Tal razonamiento sería válido si y sólo si los estados y procesos inferidos por los físicos (y los neurocientíficos) son ‘inobservables’ en el mismo sentido en que lo son los estados y procesos inferidos por los psicólogos cognoscitivos. Sin embargo, tal y como lo he mostrado aquí, ese no es el caso. Por lo tanto, la analogía de los psicólogos cognoscitivos (y de la gran mayoría de los psicólogos) no es válida.

De nuevo, algunos psicólogos cognoscitivos (e.g., Zentall, 2001) podrían rechazar el presente análisis sobre la base de una doctrina instrumentalista, arguyendo que los ‘estados’ y ‘procesos’ que postula no deberían ser vistos como literalmente reales. Bajo este argumento, tales ‘estados’ y ‘procesos’ deberían ser vistos sólo como *ficciones útiles* para generar datos experimentales de manera eficaz, sistemática y eficiente. Sin embargo, el presente artículo se concentra sobre aquellos psicólogos cognoscitivos que pretenden utilizar dichos ‘estados’ y ‘procesos’ como recursos explicativos [psicólogos que, a lo largo del presente artículo, he supuesto, siguiendo a Amundson (1985), que constituyen una porción sustancial dentro de la psicología cognoscitiva]. Estos psicólogos deben necesariamente afirmar la verdad de los enunciados que expresan relaciones entre esos estados y procesos y, en esa medida, la existencia real de los mismos. Ello debido a que, como ya lo he dicho, los estados y procesos cognoscitivos tendrán un valor explicativo sólo en la medida en que sean reales. Los ‘estados’ y ‘procesos’ cognoscitivos, vistos como meros instrumentos, no poseen valor explicativo alguno, al menos bajo los tres modelos de la explicación que he examinado a lo largo del presente artículo (modelos que son, a todas luces, realistas, en la medida en que, de nuevo, consideran al *explicans* como literalmente verdadero).

## CONCLUSIÓN: CAUSALIDAD Y DETERMINISMO EN LA CONDUCTA

Todo lo anterior hace surgir de manera especialmente patente el problema de la causalidad y el determinismo de la conducta, el cual trasciende los límites del presente artículo, en la medida en que nos adentra en cuestiones metafísicas. Sin embargo, quisiera finalizar haciendo un breve comentario al respecto.

Ciertamente, aplicaciones de los modelos ND, RE e IME (especialmente del segundo, dada la revisión de Salmon) a explicaciones que incluyen leyes neuronales en sus *explicans*, parecieran implicar que las mismas deben tomarse como leyes *causales*. Ello significaría, entre otras cosas, atribuir un papel causal en la conducta a eventos y procesos neuronales, atribución que va en contra de una de las tesis centrales del conductismo radical, a saber, el ambiente como causa principal de las respuestas (o patrones motores) que constituyen la conducta. Esta posibilidad abre toda una gama de reflexiones que no puedo examinar aquí. Sólo mencionaré tres cuestiones al respecto.

Primero, los conductistas radicales han sido notablemente ambiguos y oscuros en su uso del término 'causa'. Existen fácilmente por lo menos ocho conceptos diferentes de causa en filosofía, por lo cual habría que precisar exactamente cuál concepto estamos utilizando al afirmar que el ambiente es 'causa' de la conducta. Esta consideración, por supuesto, se aplica a todas aquellas partes del presente artículo en las cuales he utilizado el término 'causa'. Quizás, los estados y procesos neuronales jueguen un papel causal en un sentido distinto del papel causal jugado por el ambiente.

Segundo, aun cuando le atribuyamos un papel causal a los estados y procesos neuronales en la conducta, la inmensa complejidad que éstos adquieren en la conducta de un organismo total (aun en organismos relativamente simples, como ratas o palomas; ni hablemos de los humanos) puede muy bien imposibilitar la construcción de explicaciones causales efectivas. Tales explicaciones pueden ser posibles sólo en el caso de fenómenos conductuales simples que hayan sido observados en organismos simples (tales y como el que he examinado en el presente artículo a manera de ejemplo).

Finalmente, tal y como lo ha sugerido Skinner (1974), si rehusamos a atribuirles un papel causal a los estados y procesos neuronales, podríamos considerarlos como *efectos*, en lugar de cómo causas, con lo cual serían, junto con cambios en el responder, parte de *explanandum*, y no del *explanans*. Ello significaría, entre otras cosas, que dichos estados y procesos vendrían a ser parte del *fenómeno* a ser explicado. Las implicaciones de esta posibilidad son demasiado profundas y extensas como para examinarlas aquí. Basta decir, por el momento, que tal posibilidad abre el camino hacia una nueva ciencia, no de la conducta, no del sistema nervioso, sino de los *organismos* como entidades que sólo pueden ser entendidas como sistemas biopsicológicos (no sólo biológicos, no sólo psicológicos) integrados. Una ciencia no analítica, sino *sintética*. El nivel de síntesis involucrado en tal ciencia requerirá de avances conceptuales, teóricos



y metodológicos que aun están por ocurrir, quizás en un futuro lejano. Pero el camino hacia una ciencia de los organismos (la cual podríamos llamar, por el momento y a falta de un mejor término, 'organismología') ya ha sido abierto, y sólo es cuestión de tiempo, esfuerzo y dinero para que tal ciencia sea establecida como disciplina autónoma.

## REFERENCIAS

- Amundson, R. (1985). Psychology and epistemology: The place versus response controversy. *Cognition*, 20, 127-153.
- Bohm, D. y Hiley, B. J. (1993). *The undivided universe: An ontological interpretation of quantum theory*. London: Routledge.
- Carnap, R. (1966). *An introduction to the philosophy of science*. Dover: New York.
- Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Boston: Authors Cooperative.
- Derr, P., y Thompson, N. S. (1992). Reconstructing Hempelian motivational explanations. *Behavior and Philosophy*, 20, 37-45.
- Guthrie, E. R. (1933). On the nature of psychological explanations. *Psychological Review*, 40, 124-137.
- Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
- Hempel, C. G. y Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135-175.
- Lipton, P. (1991). *Inference to the best explanation*. London: Routledge.
- Mach, E. (1893/1960). *The science of mechanics: A critical and historical account of its development*. Illinois: Open Court.
- Maxwell, N. (1993). Does orthodox quantum theory undermine, or support, scientific realism? *The Philosophical Quarterly*, 43, 139-157.
- Moulines, C. U. (1982). *Exploraciones metacientíficas*. Madrid: Alianza.
- Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. Oxford: Clarendon.
- Putnam, H. (1962). What theories are not. En E. Nagel, P. Suppes y A. Tarski (Eds.), *Logic, methodology and philosophy of science* (pp. 240-251). Stanford University Press.
- Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, 74, 71-80.
- Rescorla, R. A. y Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. En A. H. Black y W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press.
- Salmon, W. C. (1998). *Causality and explanation*. Oxford University Press.
- Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science*, 3, 150-161.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193-216.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Skinner, B. F. (1977). Why I am not a cognitive psychologist. *Behaviorism*, 5, 1-10.
- Skinner, B. F. (1989). The origins of cognitive thought. *American Psychologist*, 44, 13-18.

- Stegmüller, W. (1973). *Estructura y dinámica de teorías*. Heidelberg: Springer-Verlag. Versión en español de C. U. Moulines (1983). Barcelona: Ariel.
- Suppe, F. (1977). *The structure of scientific theories*. University of Illinois Press.
- Suppe, F. (1989). *The semantic conception of theories and scientific realism*. University of Illinois Press.
- Zentall, T. R. (2001). The case for a cognitive approach to animal learning and behavior. *Behavioural Processes*, 54, 65-78.

## RESUMEN

Mi objetivo principal en el presente artículo es ofrecer un conjunto de reflexiones críticas acerca de lo que psicólogos cognoscitivos, conductistas radicales y filósofos de la ciencia han dicho acerca las explicaciones cognoscitivas. Mi punto de partida es la aplicación de tres modelos de la explicación, a saber, el modelo nomológico-deductivo, el modelo de la relevancia estadística, y el modelo de la inferencia a la mejor explicación. Según estos tres modelos, los enunciados cognoscitivos poseen poder explicativo en tanto en cuanto involucran términos teóricos que, supuestamente, son parcialmente interpretados, es decir, términos que se refieren a estados y procesos que no son directamente observables, pero que constituyen mecanismos subyacentes de la conducta. Sin embargo, un examen cuidadoso de los enunciados neurocientíficos nos obligan a reconsiderar el poder explicativo de los enunciados cognoscitivos. Mi argumento central al respecto es que aun cuando las explicaciones que involucran enunciados cognoscitivos son epistémicamente legítimas bajo los tres modelos examinados, éstos también permiten considerarlas como epistémicamente menos satisfactorias que aquellas que involucran enunciados neurocientíficos.

Palabras clave: explicación, enunciados cognoscitivos, estados y procesos cognoscitivos, inobservables, mecanismos subyacentes, dicotomía observacional-teórico, enunciados neurocientíficos, cognoscitivismo, conductismo radical.

## ABSTRACT

In the present paper, I offer a set of critical reflections on what cognitive psychologists, radical behaviorists, and philosophers of sciences have said about cognitive explanations. My starting point is an application of three models of explanation, namely, nomological-deductive, statistical relevance, and inference to the best explanation. According to the three models, cognitive statements possess an explanatory power in so far as they involve theoretical terms that supposedly are partially interpreted, that is, terms that refer to states and processes that not directly observable but constitute underlying mechanisms of behavior. However, an examination of the statements from the neurosciences forces us to reconsider the explanatory power of cognitive statements. My central argument in this regard is that although explanations of behavior that involve cognitive statements are epistemologically legitimate under the three models examined, such models also allow us to consider them as epistemologically less satisfactory than those that involve neuroscientific statements.

Key words: explanation, cognitive statements, cognitive states and processes, inobservables, underlying mechanisms, observational-theoretical dichotomy, neuroscientific statements, cognitivism, radical behaviorism.