

EPISTEMOLOGÍA E HISTORIA DE LA CIENCIA

SELECCIÓN DE TRABAJOS DE LAS XXII JORNADAS

VOLUMEN 18 (2012)

Luis Salvatico
Maximiliano Bozzoli
Luciana Pesenti
Editores



ÁREA LÓGICO-EPISTEMOLÓGICA DE LA ESCUELA DE FILOSOFÍA
CENTRO DE INVESTIGACIONES DE LA FACULTAD DE FILOSOFÍA Y HUMANIDADES
UNIVERSIDAD NACIONAL DE CÓRDOBA



Esta obra está bajo una Licencia Creative Commons atribución NoComercial-SinDerivadas 2.5 Argentina



La explicación psicológica como explicación mecanicista

*Sabrina Hamovici **

1. Introducción

Craver (2006) establece ciertos criterios para distinguir modelos explicativos de modelos no explicativos. Si se aplican estos criterios a las explicaciones ofrecidas en la psicología cognitiva, éstas resultan estar alejadas de las que se consideran “buenas” explicaciones. Aunque desde el enfoque mecanicista se concede que puedan ser útiles y cumplir un rol heurístico, no satisfacen los requisitos para constituir “buenas” explicaciones, entendidas éstas como explicaciones “completas” (Piccinini & Craver, 2011). Las explicaciones de capacidades o habilidades cognitivas parecen, en efecto, describir mecanismos. De hecho, distintos filósofos las consideran explicaciones mecanicistas (Bechtel & Wright, 2009; Glennan, 2005).ⁱ Sin embargo, no satisfacen los criterios que se aplican en las neurociencias para constituir “buenas” explicaciones. En particular, las explicaciones psicológicas no ofrecen descripciones “completas” de componentes y de sus actividades en términos físicos. Estas explicaciones no localizan directamente los componentes de los mecanismos, sino que los identifican en términos funcionales, a partir de sus roles causales. Machamer, Darden & Craver (2000) reconocen un continuo de explicaciones mecanicistas en función de cuán completa sea la descripción del mecanismo. En un extremo se encuentran los esquemas, que son descripciones abstractas e incompletas. En el otro, se encuentran las explicaciones completas, que describen en detalle todas las partes componentes del mecanismo, sus actividades y su organización. Según los autores, las explicaciones que se encuentren más cerca de los esquemas serán incompletas y, por lo tanto, su poder explicativo será muy débil. De este modo, si se evalúa a las explicaciones de la psicología cognitiva según el enfoque mecanicista, éstas se ven relegadas a un rol mínimo y provisorio.

En este trabajo sostendré que algunos modelos de explicación en psicología describen mecanismos y, si bien no satisfacen los criterios que Craver y sus colegas atribuyen a las explicaciones mecanicistas completas, cumplen cabalmente con el objetivo de explicar fenómenos psicológicos. Sostendré que, si se tienen en cuenta los objetivos que persiguen las explicaciones en psicología, los argumentos en contra del poder explicativo de sus modelos no resultan pertinentes. En la sección 2 reconstruiré los criterios que desarrollan Craver y sus colegas para evaluar las explicaciones mecanicistas, y señalaré que estos criterios resultan adecuados para las explicaciones en neurociencias. En la sección 3 reconstruiré la forma que adoptan típicamente las explicaciones psicológicas y sostendré que los criterios propuestos por Craver son demasiado fuertes para la psicología cognitiva, ya que esta última no persigue los mismos objetivos que las neurociencias. En la sección 4 intentaré mostrar cómo una explicación que no es completa puede resultar de todos modos adecuada a los objetivos de los modelos psicológicos, utilizando como ejemplo un modelo de procesamiento conceptual.ⁱⁱ

2. La explicación mecanicista en neurociencias

Las explicaciones mecanicistas muestran cómo el comportamiento de un mecanismo produce el fenómeno a explicar. Craver (2006) elabora una serie de criterios para evaluar

* UBA – CONICET, sabrinahamovici@yahoo.com.ar

cuándo un modelo de un mecanismo resulta explicativo de un fenómeno. Un mecanismo se caracteriza por estar compuesto por un conjunto de partes que realizan determinadas operaciones o actividades, y cierta organización entre las partes (e.g. temporal y espacial), que les permite producir el fenómeno que se quiere explicar. Como primer criterio, Craver sostiene que las explicaciones legítimas son aquellas que describen mecanismos. Estas se oponen a las descripciones fenoménicas, las cuales son consideradas ilegítimas porque ofrecen únicamente redescrpciones del fenómeno utilizando otros términos, pero no revelan la estructura causal o productiva que permitiría explicarlo. En síntesis, Craver opone los modelos mecanicistas (legítimos) a los modelos fenoménicos (ilegítimos). Como segundo criterio, dentro de las explicaciones legítimas, es decir, aquellas que describen mecanismos, Craver reconoce un continuo en función del detalle con el que se describen las partes y actividades del mecanismo en cada modelo. En un extremo se encuentran los meros esquemas y en el otro las explicaciones completas. Las “buenas explicaciones constitutivas” son las que detallan tanto las partes y actividades reales de un mecanismo, como su organización temporal, espacial y jerárquica (Craver 2006, p. 374).

En las explicaciones completas “ideales” no pueden figurar cajas negras, ni términos provisionales, sino que deben incluirse todos los detalles relevantes para dar cuenta del fenómeno. Así, las explicaciones mecanicistas consiguen tornar inteligibles a los fenómenos ya que, en lugar de limitarse a la redescrpción de los mismos, “involucran revelar la relación *productiva*” entre las actividades de las partes componentes y el fenómeno (Machamer, Darden & Craver 2000, p. 22). Es decir que los fenómenos se tornan inteligibles cuando se revela una estructura causal que da cuenta de su producción. Un ejemplo de explicación en la que se describe en detalle todos los factores relevantes para dar cuenta del fenómeno es la explicación de la despolarización de las neuronas en la sinapsis (Craver, 2006). Si bien esta explicación no es *idealmente* completa (ya que no incluye *todos* los factores que potencialmente podrían intervenir en el mecanismo), da una caracterización detallada del fenómeno a explicar, de las partes componentes y de sus actividades. Ofrece ciertas condiciones iniciales y finales, y, paso por paso, las actividades que permiten pasar de las condiciones iniciales a las finales, especificando la cadena causal intermedia. Así, explica el fenómeno de cambio del potencial de acción de la membrana celular durante la sinapsis como resultado de las actividades de ciertas partes componentes. Este es un ejemplo de explicación en el que no hay vacíos, ni términos provisionales, sino que todos los detalles relevantes se hallan especificados.

Si bien asumo que las explicaciones en psicología apelan a mecanismos, estos no son caracterizados en los mismos términos que los mecanismos a los que se hace referencia en las neurociencias. En general, se trata de mecanismos de procesamiento de información, cuyas partes componentes no son directamente localizables. Las partes que describen las explicaciones en psicología no se caracterizan en virtud de sus propiedades *estructurales*, sino *funcionales*. Las propiedades estructurales incluyen el tamaño, la forma, localización y orientación de los objetos. Las propiedades funcionales, en cambio, incluyen capacidades o disposiciones para realizar determinadas actividades (Piccinini & Craver, 2011). Ahora bien, siguiendo a Piccinini y Craver, los componentes de un mecanismo poseen propiedades tanto funcionales como estructurales, de modo que la descripción completa de un mecanismo debe incluir tanto las propiedades estructurales como las funcionales. Por lo general, las

explicaciones psicológicas se concentran sólo en las propiedades funcionales y dejan sin especificar las estructurales.

Si se adoptan los criterios de Craver, las explicaciones psicológicas serían legítimas, ya que describen mecanismos, puesto que dan cuenta de la estructura causal de sistemas, identificando funcionalmente sus partes componentes, así como las actividades que realizan. Sin embargo, estas explicaciones quedarían en el extremo poco útil de las explicaciones legítimas, como meros esquemas de explicación, ya que incluyen cajas negras y términos metafóricos y provisorios. Según Piccinini y Craver (2011, p. 292) los esquemas de explicación “están incompletos porque dejan afuera detalles cruciales acerca de cómo funciona el mecanismo”.ⁱⁱⁱ Las explicaciones psicológicas podrían ser consideradas explicaciones parciales, pero Craver (2006) les da una connotación casi peyorativa, señalando que aunque revelan parte de la estructura causal de un mecanismo, lo hacen utilizando partes y operaciones provisorias, que dan una ilusión de comprensión. De modo que, según los criterios de Craver, el status de las explicaciones en psicología, en tanto presentan modelos mecanicistas incompletos, es sólo el de explicaciones provisorias.

Considero importante contextualizar los dos criterios de evaluación propuestos por Craver, que distinguen, por un lado, entre modelos explicativos y fenoménicos y, por otro, entre esquemas y explicaciones completas. Estos dos criterios se proponen para evaluar las explicaciones ofrecidas en neurociencias, y, como tales, resultan adecuados para las explicaciones de esas disciplinas, que persiguen, fundamentalmente, dos objetivos. El primer objetivo es explicar diversos fenómenos relacionados con el cerebro, tales como el modo en que se desarrolla desde la infancia o cómo el sistema visual da lugar a la percepción de color. El segundo objetivo es “controlar el cerebro y el sistema nervioso” para, por ejemplo, tratar enfermedades (Craver, 2007, p. 1, mi traducción). Estos objetivos están relacionados: “Explicar el cerebro es un modo de descubrir cómo manipularlo, y manipular el cerebro es un modo de descubrir y testear explicaciones” (Craver, 2007, p. 2, mi traducción). Ahora bien, antes de evaluar las explicaciones psicológicas según los criterios mecanicistas, es preciso diferenciar sus objetivos de los de las neurociencias. Los objetivos de la psicología no incluyen principalmente el de intervenir directamente para controlar las bases físicas de los sistemas que describen (aunque hay disciplinas más ligadas a la clínica en las que esto puede hacerse, como es el caso de la psiquiatría). En la sección siguiente, analizaré las características de las explicaciones en psicología y su relación con los objetivos de esta disciplina. Esto me permitirá sostener que, dados los objetivos de la psicología, los criterios propuestos por Craver para las explicaciones en neurociencias son demasiado fuertes.

3. Las explicaciones de fenómenos psicológicos

La psicología cognitiva se propone dar cuenta de capacidades mentales específicas tales como la memoria, el razonamiento y el uso de lenguaje. El enfoque cognitivo en psicología trata al cerebro, un sistema físico, como un sistema representacional, e intenta explicar los fenómenos cognitivos postulando distintos mecanismos que procesan información. Al postular mecanismos cognitivos compuestos por representaciones y operaciones de transformación de esas representaciones, puede parecer que las explicaciones en psicología cognitiva deban responder a los dos criterios normativos, que vimos en la sección anterior, propuestos por Craver y sus colegas para las explicaciones en neurociencias. Sin embargo, si se aplicaran estos dos criterios conjuntamente, ninguna explicación psicológica resultaría

adecuada. Serían explicaciones legítimas porque satisfacen el criterio de describir mecanismos (a diferencia de los modelos fenoménicos) pero, puesto que son explicaciones incompletas, serían sólo un paso en la búsqueda de las explicaciones completas y no tendrían valor en sí mismas. Las explicaciones en psicología no identifican partes físicas de los mecanismos, sino partes funcionales, es decir, partes que se distinguen por sus papeles o roles causales^{iv}. Muchas explicaciones en psicología cognitiva caracterizan un conjunto de representaciones y de operaciones, sin describir entidades en el nivel neuronal. Por supuesto que hay algunas disciplinas dentro de las ciencias cognitivas cuyas explicaciones satisfacen los dos criterios propuestos por Craver. Por ejemplo, la neuropsicología o la neurolingüística tienen objetivos más cercanos a los de las neurociencias, ya que se proponen conectar los enfoques de procesamiento de información con sus bases neurales. En consecuencia, ofrecen explicaciones que tienen más posibilidades de satisfacer los dos criterios antes mencionados. Pero no toda la psicología cognitiva comparte estos objetivos y, por lo tanto, gran parte de las explicaciones psicológicas omite la descripción detallada de las bases neurales de los sistemas de procesamiento.

Piccinini & Craver (2011) sugieren que los análisis funcionales propuestos en psicología son explicaciones provisionarias que pueden luego completarse integrándose con explicaciones mecanicistas de sistemas neurales. Así, sostienen que, si bien son integrables en explicaciones completas y podrían cumplir un rol heurístico en la investigación, no son, en sí mismas, explicaciones *full-blown*. Glennan (2005), en cambio, defiende que las explicaciones en psicología, siendo mecanicistas, son buenas por sí mismas, sin una promesa de una mejor explicación posterior. Señala que existen métodos indirectos para evaluar mecanismos cuyas partes no pueden localizarse directamente y defiende entonces la posibilidad de presentar evidencia a favor de modelos mecanicistas, aun cuando las partes componentes no tengan una descripción física completa. Glennan (2005) muestra que se pueden diseñar experimentos alterando variables para poner a prueba distintos modelos psicológicos. Estos modelos asumen que los componentes postulados tienen una realización física, pero no pretenden manipular directamente el nivel físico. Puede defenderse que, como consecuencia, estas explicaciones no necesitan describir estructural o físicamente todas las partes componentes y sus actividades. Es decir que el criterio de Craver que diferencia, en función del detalle de las descripciones, entre explicaciones completas buenas y explicaciones incompletas meramente provisionarias, resulta demasiado fuerte para las explicaciones psicológicas. No tiene sentido exigir a las explicaciones en psicología cognitiva una descripción estructural detallada de las partes componentes si pueden elaborar y poner a prueba modelos que den cuenta de fenómenos sin apelar a este nivel de descripción. Si la estructura causal que da cuenta de la producción de los fenómenos cognitivos puede describirse en términos funcionales, estas explicaciones deberían ser adecuadas. Es decir que los criterios adoptados para evaluar las explicaciones en psicología deberían ser más moderados.

En lo que sigue analizaré un modelo psicológico de procesamiento conceptual. En términos de Craver, se trata de un esquema, ya que no describe en detalle las propiedades estructurales de los componentes del mecanismo. Intentaré mostrar que este modelo, a pesar de ser un esquema, da cuenta adecuadamente de la activación de representaciones conceptuales en ausencia de estímulos perceptivos

4. Un modelo psicológico de procesamiento conceptual

Los enfoques representacionales en psicología cognitiva asumen que los conceptos son representaciones que almacenan información sobre categorías. Los procesos de pensamiento conceptual se explican, desde estos enfoques, por medio de mecanismos que activan y manipulan las representaciones conceptuales. Barsalou y sus colegas (2003) proponen que el procesamiento conceptual está sustentado por mecanismos de reactivación de representaciones de modalidad específica, es decir, propias de cada uno de los sistemas perceptivos. Me concentraré aquí en el mecanismo de simulación o reactivación de representaciones perceptivas, que explica la activación de representaciones en ausencia de estímulos perceptivos. Según el modelo de Barsalou y sus colegas, pensar, por ejemplo, en un auto, consiste en simular o reactivar –al menos parcialmente– los estados perceptivos en los que estaría un individuo en presencia de un auto. Así, pensar consiste en reactivar representaciones perceptivas en ausencia de los estímulos externos que oportunamente las ocasionaron.

Quando percibimos visualmente un auto, se activa en el sistema visual un conjunto de detectores neurales de rasgos. Luego neuronas conjuntivas en un área de asociación cercana unen los rasgos activos y los almacenan en la memoria. Más adelante, en ausencia de *input* visual, estas neuronas conjuntivas reactivan parcialmente el conjunto original de detectores de rasgos para representar visualmente el auto (Barsalou *et al.*, 2003, p. 85, mi traducción).

Este modelo propone que pensar consiste en simular o reactivar estados perceptivos. La reactivación puede ocurrir para múltiples modalidades a la vez, ya que hay áreas de asociación o zonas de convergencia en las cuales se integra información de distintas modalidades para una categoría. Estas zonas registran patrones de activación neuronal, de modo que permiten reactivar esos patrones en ocasiones posteriores. Así, un simulador es un conjunto de representaciones de distintas modalidades para una categoría que puede ser reactivado para representar esa categoría en distintos procesos de pensamiento. Este modelo propone que, una vez que se desarrolla un simulador para una categoría, distintos subconjuntos de las representaciones del simulador pueden reactivarse para representar esa categoría en distintos contextos (Barsalou *et al.*, 2003). Barsalou y otros han realizado numerosos estudios conductuales con la intención de mostrar que las representaciones que se activan en el pensamiento son de modalidad específica. Si bien no se tiene un acceso directo a estas representaciones, y no se dispone de una caracterización detallada de las propiedades estructurales del mecanismo de simulación, los autores diseñaron estudios en los cuales se manipulan variables de los estímulos que se presentan a los individuos. Observando el desempeño de los sujetos –corrección de las respuestas y tiempos de reacción–, se pueden extraer conclusiones respecto de los mecanismos subyacentes a las tareas estudiadas.

Pecher y sus colegas (2003) se concentraron en la tarea de verificación de propiedades de un concepto. Los sujetos de este estudio debían responder si, por ejemplo, *dulce* es una propiedad de AZÚCAR. Como resultado, registraron un aumento en el tiempo de reacción para verificar propiedades de un concepto cuando había un cambio de modalidad con respecto al estímulo anterior –e.g. los sujetos respondían más velozmente que *ruidosa* es una propiedad de LICUADORA si el estímulo inmediatamente anterior había sido el par HOJAS-*crujientes* que si antes habían verificado una propiedad de una modalidad distinta,

como en el caso de ARÁNDANO-áado. Pecher *et al.* (2004) llevaron a cabo un experimento análogo en el cual los sujetos también debían verificar propiedades de un concepto. En este estudio cada concepto se presentaba dos veces, con propiedades de la misma modalidad en los dos casos o con propiedades de modalidades distintas (separados por verificaciones de propiedades de otros conceptos). Arribaron a conclusiones similares, registrando que los porcentajes de error y los tiempos de reacción eran mayores cuando la segunda propiedad correspondía a una modalidad diferente.

Estos estudios ofrecen evidencia a favor de que el pensamiento conceptual está sustentado por mecanismos de reactivación de representaciones perceptivas. A partir del aumento en el tiempo de reacción cuando se produce un cambio de modalidad, se puede inferir que el mecanismo subyacente opera con representaciones de modalidad específica. En la misma línea que defiende Glennan (2005), estos estudios sirven de ejemplo de cómo las partes componentes de un mecanismo se pueden manipular indirectamente. Si bien no se tiene un acceso directo a los formatos de las representaciones ni a las operaciones que se realizan sobre ellas, la presentación de estímulos distintos influye en el procesamiento y da cuenta de aspectos del mecanismo subyacente. Es decir que el acceso a estos mecanismos se da de manera indirecta, por medio de variaciones en el estímulo. Puesto que las respuestas obtenidas frente a estas variaciones son consistentes con las predicciones del modelo, estos estudios sirven como apoyo empírico para el modelo.

El modelo de Barsalou y sus colegas es un esquema, ya que, a pesar de mencionar aspectos estructurales —como los sistemas neuronales correspondientes a la percepción—, no están detalladas en términos físicos las actividades que llevan a almacenar información y a reactivarla en ausencia de los estímulos perceptivos. El modelo no localiza directamente cada concepto en el cerebro ni reconstruye en términos físico-químicos los procesos de almacenamiento y reactivación de conceptos. Tiene términos provisionales y cajas negras y no detalla todos los pasos que llevan desde un estado inicial a un estado final. Craver concedería que no se trata de una mera redescipción, ya que presenta una estructura causal que da cuenta del fenómeno, es decir que concedería que no es un modelo fenoménico, sino mecanicista. Pero según los criterios mecanicistas sería sólo un esquema incompleto y, por lo tanto, provisional. Sin embargo, este modelo cumple con el objetivo de explicar el fenómeno que se propone, puesto que vuelve inteligible un aspecto del procesamiento conceptual por medio de la descripción de un mecanismo que permite almacenar y reactivar representaciones conceptuales. Aunque no ofrezca descripciones estructurales detalladas, el modelo apela a las partes componentes del mecanismo de simulación, así como a sus actividades y organización, para explicar los procesos de almacenamiento y posterior recuperación de información sobre categorías. Este modelo resulta adecuado, porque se ajusta a los objetivos de la psicología cognitiva, dando cuenta de una capacidad cognitiva superior, además de poder cumplir con otros criterios de adecuación, como parsimonia o coherencia con el resto de nuestro conocimiento. Por otra parte, es un modelo que se ha puesto a prueba y cuenta con apoyo empírico. Si bien en los estudios mencionados no se identifican ni manipulan directamente las partes componentes del mecanismo, se tiene un acceso indirecto a las mismas por medio de variaciones en los estímulos. Así, este modelo es un ejemplo de cómo las explicaciones psicológicas consiguen hacer inteligibles ciertos fenómenos, aunque no ofrezcan modelos mecanicistas completos.

5. Conclusiones

En este trabajo intenté mostrar que, si bien las explicaciones psicológicas no detallan todas las propiedades estructurales de los componentes de mecanismos, sino que describen principalmente partes funcionales, pueden tornar a los fenómenos inteligibles, y pueden resultar explicativamente adecuadas. Sostuve que las explicaciones psicológicas describen mecanismos, puesto que identifican partes componentes y sus operaciones, así como la organización entre las mismas que da cuenta de la realización de los fenómenos psicológicos en cuestión. El hecho de que las partes componentes se identifiquen de manera funcional, indirectamente, no vuelve a las explicaciones menos adecuadas. Es necesario distinguir entre los objetivos de las neurociencias y los de la psicología, antes de evaluar las explicaciones de la última en función de los objetivos de las primeras.

Notas

i Adopto la noción mecanicista de explicación y no es mi intención discutir distintas nociones de explicación.

ii No me ocuparé de presentar el debate general sobre la explicación psicológica, ya que ello llevaría gran parte del trabajo. Mi objetivo es desarrollar directamente la crítica a la aplicación de los criterios de adecuación mecanicistas a las explicaciones psicológicas.

iii Mi traducción.

iv Esto no quiere decir que nieguen que los componentes tengan una realización física, sino que los identifican por sus características funcionales.

v Utilizo la nomenclatura habitual de mayúsculas para los nombres de los conceptos

Bibliografía

- BARSALOU, Lawrence, SIMMONS, Kyle; BARBEY, Aaron; WILSON, Christine. Grounding conceptual knowledge in modality-specific systems, *TRENDS in Cognitive Science* 7 (2): 84-91, 2003.
- BECHTEL, William; WRIGHT, Cory. What is psychological explanation?, Pp: 113-130, en CALVO Paco & SYMONS, John (eds.). *Routledge Companion to the Philosophy of Psychology*. London: Routledge, 2009.
- CRAVER, Carl. When mechanistic models explain. *Synthese* 153 (3): 355-376, 2006.
- CRAVER, Carl. *Explaining the Brain*. Oxford: Clarendon Press, 2007.
- GLENNAN, Stuart. Modeling mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 443-464, 2005.
- MACHAMER, Peter; DARDEN, Lindley; CRAVER, Carl. Thinking about mechanisms. *Philosophy of science* 67 (1): 1-25, 2000.
- PECHER, Diane, ZEELENBERG, René; BARSALOU, Lawrence. Verifying different-modality properties for concepts produces switching costs. *Psychological Science* 14 (2): 119-124, 2003.
- PECHER, Diane; ZEELENBERG, René, BARSALOU, Lawrence. Sensorimotor simulations underlie conceptual representations. Modality-specific effects of prior activation. *Psychonomic Bulletin & Review* 11 (1): 164-167, 2004.
- PICCININI, Gualtiero, CRAVER, Carl. Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese* 183 (3): 283-311, 2011