

Filosofía de la mente y Psiquiatría: alcances y límites de una perspectiva naturalista para el estudio de los delirios*

Recibido: noviembre 24 de 2016 | Aceptado: junio 15 de 2017
DOI: 10.17230/co-herencia.14.27.7

Emilia Vilatta**

emiliavilatta@gmail.com

Resumen En el presente artículo me concentraré en el estudio filosófico de los delirios, como un caso ejemplificador del vínculo que pueden establecer la filosofía de la mente y la psiquiatría. Frente a versiones radicalmente naturalistas, que proponen variantes reduccionistas o eliminativistas para la explicación de ciertos fenómenos mentales y sus variantes “anormales”, defenderé una versión *moderada* de la perspectiva naturalista. Al respecto, señalaré que retener algún grado de simpatía hacia el naturalismo en las investigaciones filosóficas sobre los delirios es necesario para desarrollar teorías sobre las creencias empíricamente informadas que no resulten contradictorias con los desarrollos actuales en psicología cognitiva y neurociencias. No obstante, afirmaré también que una perspectiva exclusivamente naturalista -tanto por parte de la filosofía como de la psiquiatría misma- es incapaz de dar cuenta del contexto normativo en donde juegan las creencias. Los criterios normativos, especialmente externos (sociales y pragmáticos) que se emplean para clasificar ciertas creencias como delirantes no resultan inteligibles bajo una explicación puramente natural. Por el contrario, para comprender las creencias delirantes *qua* delirantes, será preciso acercarnos a un abordaje híbrido que pueda contemplar las causas naturales del fenómeno así como su evaluación normativa.

Palabras claves:

Psiquiatría, filosofía de la mente, naturalismo, trastornos psiquiátricos, delirios, reduccionismo, eliminativismo.

Philosophy of Mind and Psychiatry: Scope and Limits of a Naturalistic Perspective for the Study of Delusions

This paper focuses on the philosophical study of delusions as an example of the link that can be established between the philosophy of mind and psychiatry. Against radically naturalistic versions that suggest reductionist or eliminative variables to explain certain mental phenomena and their ‘abnormal’ variables, I will stand up for a *toned-down* version of the naturalistic perspective. In this sense, I will point out that it is necessary to keep some level of sympathy towards naturalism in the philosophical research on delusions to develop the theories about the empirically informed beliefs that do not contradict

* Este artículo presenta resultados parciales de la investigación de doctorado, financiada por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el Instituto de Humanidades (IDH) de la Universidad Nacional de Córdoba, Argentina. A su vez, se inscribe en la línea de investigación de los siguientes proyectos: (i) “Intencionalidad: atribución y contenido”, subsidiado por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Córdoba (SECYT-UNC) y dirigido por las Doctoras Carolina Scotto y Laura Danón. (ii) “La naturaleza y estructura de los contenidos mentales”, subsidiado por el Fondo para la Investigación Científica y Tecnológica (FON-CyT) y dirigido por el Dr. Daniel Kalpokas. Y (iii) “Epistemología de la Biología y la Psiquiatría”, subsidiado por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Córdoba (SECYT-UNC) y dirigido por la Dra. Mariana Cruz y el Dr. Carlos Arias.

the current developments on cognitive psychology and neurosciences. Nevertheless, I will also state that an exclusively naturalist perspective –from both philosophy and psychiatry itself– cannot render an account of the regulatory context within which beliefs develop. Normative, particularly external, criteria (social and pragmatic) used to classify certain beliefs as delirious are not intelligible under a purely natural explanation. Conversely, to understand delusional beliefs as delusional, it will be necessary to take an interest in a hybrid approach that considers the natural causes of the phenomenon as well as its normative assessment.

** Miembro del Consejo Nacional de Investigaciones Científicas y Técnicas - Instituto de Humanidades - Universidad Nacional de Córdoba (CONICET-IDH-UNC). Doctoranda en Filosofía por la Universidad Nacional de Córdoba, Argentina.

Key words:

Psychiatry, philosophy of mind, naturalism, psychiatric disorders, delusions, reductionism, eliminativism.

La filosofía de la psiquiatría es un área de investigación interdisciplinaria interesada en diversos problemas conceptuales y empíricos que surgen de la observación de fenómenos psicopatológicos. Durante los últimos diez años, este campo de estudio ha experimentado un renovado interés a la luz del trabajo conjunto de psiquiatras, psicólogos y filósofos. Se ha gestado en particular un vínculo fructífero entre la filosofía de la mente y la psiquiatría, bajo el convencimiento de que esta área de la filosofía podría brindar herramientas conceptuales a la psiquiatría para la comprensión de los trastornos mentales, así como la psiquiatría podría nutrir con aportes teóricos y evidencia empírica a la filosofía de la mente. Como resultado del trabajo conjunto sobre cuestiones empíricas y conceptuales, no solo se ha logrado un avance en materia de discusiones teóricas, sino que también se han materializado diversas aplicaciones psicoterapéuticas en áreas como la psiquiatría y la psicología clínica.

Uno de los problemas claves que ha sido objeto de interés recurrente es la naturaleza y explicación de los trastornos psiquiátricos, es decir, si los trastornos psiquiátricos son *mentales* y, por lo tanto, deben definirse y explicarse en términos psicológicos intencionales; o si, más bien, son trastornos *cerebrales* y deben ser definidos y explicados en términos cerebrales (neurofisiológicos o incluso moleculares). Esta última forma de entender los trastornos psiquiátricos ha cobrado impulso en los últimos años, debido al acelerado avance de las neurociencias, que ha revitalizado la denominada “psiquiatría biológica”.

Esta corriente de la psiquiatría está realizando actualmente numerosas investigaciones bajo el supuesto de que los trastornos mentales son cerebrales (Andreasen, 2001; Charney, Nestler & Bunney, 1999; Kendler, 2012). Dichas investigaciones implican diversas líneas de trabajo, entre ellas: el estudio mediante técnicas de neuroimágenes, el análisis de la conectividad de las redes cerebrales, la investigación de posibles variantes genéticas que podrían llevar a determinadas anormalidades cerebrales, etcétera. Desde esta perspectiva se asume que detrás de cada categoría diagnóstica existe un proceso fisiopatológico en el sistema nervioso, el cual es atribuido hipotéticamente a una conjunción de factores genéticos y ambientales, que generarían una alteración en la fisiología normal de alguna estructura del sistema nervioso. Según este modelo, cuando un proceso tal se instaura, da lugar a una serie de alteraciones cognitivas y comportamentales que son consideradas la expresión sintomática de la fisiopatología subyacente (Bracken, Thomas, Timimi & Yeomans, 2012; Read, Mosher & Bentall, 2015). De este enfoque ha surgido la promesa de que la investigación neurobiológica, a medida que incorpore nuevas tecnologías, será capaz de descifrar los mecanismos fisiopatológicos específicos de los distintos trastornos psiquiátricos y generar pruebas de laboratorio para identificar sus marcadores biológicos (Bracken *et al.*, 2012). Dado que este enfoque considera que los trastornos psiquiátricos son cerebrales, sus defensores han argumentado que la noción de trastorno mental debe ser redesignada en términos cerebrales (Andreasen, 2001).

Frente a este paradigma tecnológico-biologicista, que pretende edificarse sobre evidencia científica, en las últimas décadas han aparecido sectores críticos, tanto desde el ámbito filosófico como desde la psiquiatría misma. La principal crítica surge precisamente del análisis cuidadoso de la evidencia científica existente, pues se ha encontrado que en numerosos casos dicha evidencia no apoya los supuestos del modelo (Timimi, 2014; Chaufan & Joseph, 2013). En particular, respecto a la etiología no ha sido posible encontrar marcadores biológicos que tengan utilidad clínica para ningún trastorno psiquiátrico (Bracken *et al.*, 2012; Chaufan & Joseph, 2013; Moncrieff & Middleton, 2015; Timimi, 2014).

Debido a la acumulación de evidencia que entra en conflicto con el paradigma tecnológico-biologicista dominante, en los últimos años numerosos autores han reclamado la necesidad de un cambio de perspectiva para abordar los problemas psiquiátricos, aunque no todos apoyan un cambio en el mismo sentido. Algunos autores argumentan que la psiquiatría debería adoptar una identidad aún más tecnológica y biomédica (neuropsiquiatría) y ligarse definitivamente a la investigación en neurociencia, genética y farmacología (Kapur, Phillips & Insel, 2012). Otros autores arguyen que esta postura no resuelve los problemas del modelo actual, ya que asumir que la práctica clínica debe realizarse en el marco de la neuropsiquiatría sitúa el foco del problema en el cerebro de las personas diagnosticadas, un cerebro descontextualizado del resto del organismo y, fundamentalmente, de su entorno. Estos autores enfatizan que los problemas de salud mental involucran dimensiones culturales, sociales y psicológicas, que no pueden ser captadas por la epistemología de la biomedicina (Bracken *et al.*, 2012; Chaufan & Joseph, 2013).

A la vez, en el marco de este grupo disidente, algunos autores del ámbito de la filosofía de la mente han reaccionado desarrollando producciones alternativas que pretenden preservar la autonomía del carácter de lo mental en los trastornos psiquiátricos; así, han señalado que estos deben ser entendidos en términos de trastornos de la conciencia, la intencionalidad y la racionalidad, y, en consecuencia, deben ser explicados en un nivel personal con un vocabulario intencional. Con el fin de discutir algunas cuestiones en torno al estudio de los delirios, me interesa retomar aquí en particular la “*tesis del cerebro intacto*”, propuesta por G. Graham (2010, 2013), para defender una versión *moderada* de la perspectiva naturalista en la filosofía de la mente y la psiquiatría.

En lo que sigue presentaré algunos de los supuestos que subyacen al enfoque radicalmente naturalista y biologicista, que propone variantes reduccionistas o eliminativistas para explicar ciertos fenómenos mentales y sus variantes “anormales”. Luego, retomaré y profundizaré las críticas realizadas por Graham y su distinción entre trastornos *en* el cerebro y trastornos *del* cerebro. En la última sección presentaré algunas consideraciones a favor de una perspectiva híbrida para el estudio de los trastornos psiquiátricos en general y de los delirios en particular.

Reconociendo los méritos de la psiquiatría biológica, pero a la vez sus limitaciones, podremos abogar por un enfoque naturalista *moderado*. Esto implicará retener cierta simpatía hacia el naturalismo para desarrollar teorías sobre las creencias (ordinarias y delirantes) empíricamente informadas, que no resulten contradictorias con los desarrollos actuales en psicología cognitiva y neurociencias; pero a la vez habrá que conservar el carácter normativo de lo mental, que nos permite comprender y evaluar socialmente las creencias delirantes *qua* delirantes.

Algunos de los supuestos de la (Neuro)psiquiatría biológica

Algunas de las razones detrás de la concepción de los trastornos psiquiátricos defendida por la psiquiatría biológica están representadas en un argumento central que podría esquematizarse del siguiente modo:

- 1) Lo mental, de alguna manera, no es más que lo físico, lo neuronal.
- 2) X es un trastorno mental.
- 3) Por lo tanto, X es un desorden físico, un trastorno del cerebro.

Esto ha llevado, como señala Graham (2010), a un argumento similar:

(i) El concepto mismo de trastorno mental es un anacronismo histórico y científico. Presupone que a diferencia de los trastornos somáticos, los trastornos mentales son trastornos de una cosa no-física, ya sea sustancia o propiedad, denominada “mente”.

(ii) No hay cosas que no sean físicas. La suposición de que las mentes inmateriales existen simplemente no se puede combinar con una visión médico-científica del mundo.

(iii) Por lo tanto, cualquier trastorno llamado “mental”, si es que es un trastorno de algún tipo, tiene que ser de naturaleza física. Calificar un trastorno como “mental” es una equivocación metafísica.

Por otra parte, como Graham (2010) ha notado respecto a la conclusión de este último argumento, la misma es usualmente extendida del siguiente modo: (iii)* Los trastornos mentales son físi-

cos. Que sean trastornos físicos significa que son trastornos *del* cerebro, es decir, del sistema nervioso. A partir de estos supuestos, los defensores de este enfoque afirman que (i) los trastornos mentales son un subtipo de trastornos cerebrales, (ii) que los trastornos mentales son enfermedades¹ y (iii) que para una comprensión acabada de los mismos, deben ser descriptos en términos de un mal funcionamiento cerebral, recurriendo al lenguaje de las neurociencias (Andreasen, 1984, 2001, 2005). Esto es así porque bajo la concepción de la psiquiatría biológica se asumen posiciones fisicalistas (para algunos de sus defensores los estados mentales se identifican con estados cerebrales, *i.e.*, materialismo de la identidad, mientras que otros toman posiciones eliminativistas y consideran que las descripciones de los trastornos eventualmente serán realizadas solo en términos neuronales o moleculares).

Asimismo, los defensores de la Psiquiatría Biológica suelen emplear un argumento similar referido a que no hay nada distintivo en los trastornos mentales respecto a los trastornos físicos/cerebrales.² Este argumento tiene la siguiente forma:

(i) Las enfermedades somáticas (físicas) son un tipo de trastorno, y las enfermedades mentales son otro tipo de trastorno. Cada uno es una especie distinta de trastorno.

(ii) Las enfermedades somáticas son enfermedades en el cuerpo. Para que un trastorno mental sea un tipo de trastorno distinto de las enfermedades físicas, debe ser un trastorno en una entidad no-física: la mente.

(iii) Sin embargo, ninguna cosa o entidad es no-física.

(iv) Por lo tanto, no existen (de manera distintiva) los trastornos mentales.

¹ Que los trastornos mentales sean enfermedades, tal como lo son las enfermedades somáticas, es aún discutido (Szasz, 2008). Graham (2010) en particular señala que tratar los trastornos mentales como enfermedades probablemente no sea el mejor modo de entenderlos, puesto que no cumplen muchos de los requisitos para ser considerados como tales: no tienen un número limitado de causas, como las enfermedades a menudo tienen, y carecen de márgenes relativamente precisos que permitan diferenciar los casos normales de los patológicos.

² En particular, este argumento es defendido por Nancy Andreasen (1984, 2001, 2005), una de las más prominentes defensoras de la psiquiatría biológica.

No obstante, las conclusiones de estos argumentos son discutibles, en particular del (iii)* y el (iv). El hecho de que nos comprometamos con una visión naturalista sobre el origen de la mente no implica *per se* la asunción de una posición reduccionista o eliminativista. Aunque asumamos como posición ontológica un materialismo de la identidad, no nos vemos obligados a explicar los trastornos psiquiátricos en términos de procesos cerebrales. La primera cuestión es de orden ontológico, mientras que la segunda es de orden epistemológico. Antes de proseguir, nos detendremos a considerar algunas cuestiones relativas al estudio de los delirios.

El caso de los delirios

En el marco de este debate, un fenómeno que ha atraído especialmente la atención de los filósofos ha sido el de los delirios, ya que entendidos como casos paradigmáticos de irracionalidad, suponen un desafío significativo para buena parte de las teorías filosóficas clásicas sobre la explicación e interpretación de la conducta, las cuales suelen comprometerse con estándares de racionalidad muy elevados. Esto pondría en jaque la posibilidad de explicar en términos intencionales los delirios. Al respecto, algunos autores (Blackwood, Howard, Bentall & Murray, 2001; Coltheart, 2007; Stone & Young, 1998) han señalado, en línea con la psiquiatría biológica, que al igual que el resto de los trastornos psiquiátricos, los delirios deben ser explicados en términos de un mal funcionamiento cerebral. Antes de continuar, cabe repasar cómo son usualmente conceptualizados los delirios.

Los delirios han sido considerados como el sello patognomónico de la locura, siendo un síntoma crucial de la psicosis (en particular de la esquizofrenia) y una característica prominente de otras condiciones psiquiátricas y neurológicas. Si bien hay mucha controversia respecto a cómo definirlos, en particular porque ninguna de las definiciones actuales parece ofrecer criterios suficientes y necesarios (Miyazono & Bortolotti, 2015), la definición provista por la quinta edición del Manual diagnóstico y estadístico de los trastornos mentales (DSM V) dice lo siguiente:

Creencia falsa basada en deducciones incorrectas sobre la realidad externa que se mantiene firmemente a pesar de lo que el resto de las personas creen, y a pesar de las pruebas evidentes e indiscutibles de lo contrario. La creencia no es ordinariamente aceptada por otros miembros de la cultura o subcultura de la persona (esto es, no es un principio de la fe religiosa). Cuando una falsa creencia implica un juicio de valor, se considera como un delirio solo cuando el juicio es tan extremo como para desafiar la credibilidad. (American Psychiatric Association, 2013, p. 824)

Por otra parte, en el marco de las discusiones filosóficas sobre los delirios, los defensores del enfoque doxástico, esto es, delirio como creencia (Bayne & Pacherie, 2005; Bortolotti, 2005; Miyazono & Bortolotti, 2015), han señalado que hay al menos tres componentes centrales de la creencia delirante:

1. Es formada sin suficiente o adecuada justificación.
2. Está compartimentada/circunscripta.³
3. Es mantenida frente a fuerte evidencia contraria.

Ahora bien, aceptando, no sin algunas controversias, que los delirios son estados doxásticos, es decir, creencias, la pregunta en cuestión es en qué términos explicar estas creencias “anormales” y sus características particulares. ¿Es posible dar cuenta de ello apelando a explicaciones de nivel puramente subpersonal, como pretenden los defensores de la psiquiatría biológica? A continuación, recuperando la *tesis del cerebro intacto*, de Graham (2010, 2013), esbozaremos algunas razones para responder negativamente a esta pregunta y defender un abordaje híbrido de las explicaciones sobre los delirios.

³ En la literatura se mencionan tres tipos de circunscripción: (i) inferencial, por cuanto los sujetos podrían fallar en esbozar las consecuencias lógicas de sus creencias delirantes y a veces no lograrían resolver las contradicciones entre el delirio y el resto de sus creencias; (ii) conductual, ya que en algunos casos los sujetos delirantes no se comportan como si creyeran en el contenido de sus delirios; (iii) afectiva, pues se afirma que en algunos casos los sujetos tampoco exhiben las respuestas afectivas apropiadas a sus delirios en circunstancias relevantes. Cabe señalar, no obstante, que estos tipos de circunscripción no se hayan presentes en todos los casos de delirios.

Normatividad, función y éxito explicativo

Como hemos señalado previamente, incluso si aceptamos una explicación *natural* sobre el origen de la mente y adscribimos a algún tipo de posición fisicalista, de esto no se sigue necesariamente que debamos explicar los trastornos psiquiátricos solo en términos físicos/causales (ya sea en el vocabulario de la neurociencia de redes, moléculas o genes). La mente podría ser algo físico, es decir que los procesos mentales podrían ser simplemente cerebrales; sin embargo, algunas formas de comportamiento -como las asociadas a trastornos psiquiátricos- pueden comprenderse mejor, al menos provisoriamente, haciendo referencia a estados intencionales. Esta perspectiva no remite ya a una cuestión ontológica sobre la mente y las relaciones mente-cerebro, sino a un asunto explicativo en búsqueda de obtener el mejor modo de comprender los trastornos psiquiátricos.

Adicionalmente, aunque uno mantenga cierto tipo de *optimismo epistemológico*, como parecen asumir los defensores de la psiquiatría biológica,⁴ al menos por el momento las neurociencias no pueden proveernos de una explicación satisfactoria sobre la constitución física de los estados mentales. Por supuesto, a pesar de que las neurociencias no nos ofrezcan tales explicaciones, aún puede ser verdadero el fisicalismo respecto a la ontología de los estados mentales, y una explicación en términos de causas *naturales* podría, eventualmente, ser encontrada. Al respecto, Graham (2010) ha propuesto como solución de compromiso no adoptar una postura metafísica particular sobre la mente, inclinándose a un “*agnosticismo metafísico*”. Bajo esta posición, nuestra capacidad de explicar y comprender un trastorno mental no tiene relación directa con la verdad o falsedad del fisicalismo como tal, siempre que asumamos que los estados mentales pueden ser vehículos causalmente eficaces de contenido intencional.

⁴ Como respuesta a la falta de consistencia en el campo de la psiquiatría biológica acerca de los resultados de la búsqueda de la etiología de los trastornos, los defensores de este enfoque suelen aludir a una incapacidad epistémica *provisoria* de la disciplina (y de otras disciplinas adyacentes, como las neurociencias), que eventualmente será superada, cuando se logren los avances tecnológicos necesarios junto a una mejor definición de los constructos utilizados (Balestri, Calati & de Ronchi, 2014). Por el contrario, los detractores de este enfoque discuten esta posición y señalan que la falta de hallazgos significativos no se debe a una limitación tecnológica o metodológica, sino a presuposiciones erróneas sobre las presuntas causas de los fenómenos (Chaufan & Joseph, 2013; Timimi, 2014).

Según mencionamos con anterioridad, actualmente muchos filósofos y psiquiatras consideran que los trastornos mentales, por el solo hecho de concebir que tienen una base neurofisiológica, son trastornos *del* cerebro. Así, los defensores de la psiquiatría biológica admiten que existen los trastornos “mentales”, y en ese sentido son realistas, pero niegan que sea útil clasificarlos como inequívocamente mentales. Consideran que, más que trastornos de la mente, son un subtipo de trastornos del cerebro, y no encuentran razones para la distinción entre trastornos mentales y cerebrales.

En oposición a esta concepción, Graham (2010) ha propuesto la *tesis del cerebro intacto*, la cual sostiene que aunque una persona tenga un trastorno mental, su cerebro puede estar sano de acuerdo con estándares neurológicos. En función de esta postura, podemos distinguir entre los trastornos que se basan *en* el cerebro (los estrictamente mentales, por ejemplo: una depresión, un trastorno delirante o una fobia) de los trastornos *del* cerebro (una lesión cerebral, mal de Alzheimer, mal de Parkinson, etcétera). En palabras de Graham: “los trastornos mentales pueden ser trastornos *en* el cerebro (realizados en el cerebro o ‘hardware’) sin ser trastornos *del* cerebro (sin un mal funcionamiento neuronal o una lesión)” (2010, p. 24).

La razón principal aducida para promover esta distinción es que los *criterios normativos* para evaluar el funcionamiento adecuado del cerebro son diferentes de aquellos usados para evaluar un comportamiento racional; tal como, según una analogía frecuentemente utilizada, los estándares de un buen hardware no son los mismos que los de un buen software. En consonancia con esta distinción, algunos autores señalan:

Si los equipos son completamente físicos, y sin embargo se pueden distinguir de manera significativa problemas de hardware y de software, entonces es posible que aunque los seres humanos sean completamente bioquímicos todavía se pueda diferenciar los estados mentales y sus problemas, de los estados no mentales. (Arpaly, 2005, p. 48)

La distinción propuesta por Graham (2010) puede tener repercusiones no solo conceptuales, sino también prácticas. Por ejemplo, en la discusión epistemológica referida a si estamos dispuestos o no a

dejar que la psiquiatría sea reducida a las neurociencias; o en la elección del tipo de tratamiento adecuado para personas con trastornos psiquiátricos. Si operase una distinción entre trastornos cerebrales y mentales, los primeros (como el Alzheimer o el Parkinson) podrían ser más susceptibles de tratamiento médico-farmacológico, mientras que los segundos podrían ser tratados exitosamente mediante intervenciones psicoterapéuticas, como tratamiento adicional o único, dependiendo del caso. Asimismo, la distinción de Graham entre los trastornos *en y del* cerebro resulta metodológicamente interesante, ya que al desconectar la conceptualización y explicación de los trastornos de la tesis ontológica sobre qué son en última instancia, se preserva una perspectiva teórica especial para entender qué hace que un trastorno sea mental.

Por otra parte, consideremos a continuación razones adicionales para resistir al reduccionismo de la psiquiatría biológica. En primer lugar, hallamos ciertas dificultades relativas a la determinación del buen o mal funcionamiento cerebral; en segundo, encontramos inconvenientes referidos al éxito actual de las explicaciones neurocientíficas.

Respecto al primer punto, cabe señalar que quienes afirman que los trastornos mentales son un subtipo de trastorno cerebral parecen asumir que los mecanismos neuronales involucrados en cada capacidad psicológica han sido creados para servir a la salud mental. Sin embargo, la evolución no ha diseñado el cerebro de modo que cada una de sus fallas en la conducta saludable desde el punto de vista psiquiátrico/psicológico correspondan a factores no saludables desde una perspectiva neurológica. De hecho, algunas investigaciones señalan que el cerebro no siempre actúa de la mejor manera, aunque no esté neurológicamente dañado y, en consecuencia, “parece existir una gran brecha entre lo que los cerebros pueden hacer y lo que es mejor para los seres humanos” (Stich, 1990, p. 154). El cerebro, por ejemplo, debido a su plasticidad, nos permite aprender de un modo variable (y a través de reforzamientos), lo cual nos posibilita adaptarnos a diferentes ambientes que pueden ser desconocidos o impredecibles. Pero esta capacidad de aprendizaje, con la que hemos sido diseñados, es compatible con el hecho de que una persona caiga en un patrón de aprendizaje de conductas adictivas, que lo lleve a

un mal funcionamiento psicológico. Vemos así que una conducta psicológica poco sana o deteriorada no representa *ipso facto* una falla del cerebro para realizar sus funciones.

A esto se añade el problema de determinar cuáles son las funciones del cerebro (o de las áreas o centros cerebrales específicos), pues, por una parte, aún no poseemos siquiera un mapa completo del funcionamiento “normal” del cerebro⁵ y, por otra, todavía está abierto el debate -en particular en filosofía de la biología- sobre cómo determinar cuál es la función normal o apropiada de un órgano y, en consecuencia, un mal funcionamiento. En relación con este punto, McKey y Dennett (2009) han sugerido un ejemplo interesante para mostrar las dificultades de establecer cuál es la función adecuada de un órgano o sistema. Pensemos en el sistema inmunológico y los cientos de casos en los cuales dicho sistema “rechaza” un órgano en un trasplante que podría salvar una vida. En la medida en que la función del sistema es atacar cuerpos extraños, se han obtenido buenos resultados. Sin embargo, cuando interpretamos que el sistema tiene la función más general de proteger la salud corporal, se vuelve menos claro si ha funcionado normalmente. En palabras de McKey y Dennett (2009, p. 495): “éste es el problema de las intenciones evolutivas impuestas por un tercero. Quizás todo depende de qué función elegimos imponer al sistema. ¿Está toda la funcionalidad sólo en el ojo del espectador?”. Estas dificultades presentes en el caso de sistemas de los que se tiene actualmente un gran conocimiento se exacerban cuando se trata de la asignación de funciones al cerebro, pues, como recién señalamos, aún no resulta claro siquiera el modo en que funciona en sujetos sin trastornos.

En cuanto al segundo punto, para justificar la clasificación de los trastornos mentales como un subtipo de trastornos cerebrales, no basta con que el fisicalismo sea cierto; además, la descripción de las causas de tales trastornos debe poder realizarse de manera comple-

⁵ Desde el año 2010 se está desarrollando el proyecto *Conectoma Humano (Human Connectome Project)*, el cual tiene como propósito lograr un mapeo integral del cerebro (en analogía con el proyecto *Genoma*) para obtener datos sobre la organización de las conexiones estructurales y la conformación de las dinámicas funcionales. Si bien se han realizado importantes hallazgos en cuanto al cerebro humano en estadios prenatales (Miller, Ding & Sunkin, 2014), las investigaciones aún se encuentran muy lejos de poder establecer completamente el conectoma adulto.

ta en términos de descripciones neurocientíficas, prescindiendo del vocabulario mentalista. Según señalamos, de acuerdo con quienes defienden la idea de que los trastornos mentales son cerebrales, detrás de cada categoría diagnóstica existe un proceso fisiopatológico en el sistema nervioso y, en consecuencia, la categorización de una afección como un trastorno psiquiátrico debería hacerse en términos de las estructuras neuronales que se consideran causalmente responsables de los síntomas. Es decir que el trastorno psiquiátrico tendría que reflejar siempre alguna neuropatología distintiva: algún área del sistema neuronal se debería encontrar relevantemente afectada por una determinada lesión o alteración significativa. Por lo tanto, las presuntas causas que se busca identificar también deben referirse a daños cerebrales/neuronales, y ningún vocabulario psicológico intencional tiene que ser necesario para reconocer las causas de los síntomas de los trastornos del cerebro.

Cuando los neurocientíficos investigan el cerebro y sus afecciones (en el caso de la psiquiatría biológica) se rigen por un principio denominado “*principio de la suficiencia causal explicativa del dominio físico/neuronal dañado*” (Graham, 2013, p. 519). Este sostiene lo siguiente: si una afección neuronal es un trastorno del cerebro, debe tener una o varias causas físicas/neuronales que admitan descripciones físicas/neuronales, y dichas causas deben ser en sí mismas afecciones que admitan la clasificación de dañado o trastornado. Entonces, si los trastornos mentales constituyen un subtipo de trastornos cerebrales, la categoría de trastorno mental debería obedecer al mismo principio.

Dicho principio se mantiene firme en la neurociencia actual y en campos médicos relacionados, a pesar de los reiterados fracasos en encontrar algún criterio común aceptable para saber qué tipo de afecciones o estados del cerebro calificarían como disfuncionales o patológicos. Asimismo, los estándares para satisfacer este principio no están establecidos de modo fijo ni son el producto de un consenso entre científicos, basado en evidencias. Más bien, en la práctica, “la utilización de los conceptos cerebrales aplicados a los trastornos, tiende a depender de juicios e interpretaciones, algunas veces controversiales, sobre las funciones de los componentes del cerebro” (Graham, 2013, p. 13).

Hasta aquí podemos decir que si las causas de los trastornos no pueden ser descriptas enteramente en términos neurocientíficos, entonces los trastornos mentales no deberían ser clasificados como un subtipo de los trastornos cerebrales. El mero apoyo ontológico del fisicalismo sobre lo mental no es suficiente para garantizar que los trastornos mentales sean trastornos del cerebro. También es necesario poder elaborar explicaciones exitosas de tales trastornos en términos exclusivamente neurocientíficos. El fisicalismo, en caso de que sea correcto, asegura simplemente que los trastornos tienen una base *en* el cerebro, pero solamente una comprensión explicativa acabada de la constitución de un trastorno en términos neurocientíficos permite decir que un trastorno sea *del* cerebro. De este modo, siguiendo el espíritu de la tesis de Graham, podemos evitar relacionar los méritos del fisicalismo con las vicisitudes empíricas de la explicación, distinguiendo, entre los argumentos a favor del fisicalismo, la tesis ontológica, que asegura que todo lo que existe es materia física, de la tesis lingüística-explicativa, que sostiene que todo lo físico puede ser capturado y descripto en el lenguaje y los conceptos de la ciencia física.

De este modo, si el vocabulario neurocientífico no resulta suficiente para explicar los trastornos psiquiátricos, es válido apelar al vocabulario intencional, que sin descartar el primero, puede complementarlo en explicaciones híbridas.⁶ Esto permitiría tener la libertad conceptual necesaria para sostener que un trastorno mental no es cerebral (*del* cerebro), haciendo uso de conceptos y evidencias neurocientíficas, pero a la vez de conceptos psicológicos (para los fines explicativos), aun cuando se apoye un fisicalismo (como tesis ontológica) respecto a las entidades a las cuales nos referimos con este último tipo de conceptos. Entonces, si se necesita hacer referencia a las creencias, deseos o experiencias conscientes de un agente para entender aspectos críticos relevantes de un trastorno -de sus causas e influencias en la conducta-, este califica como *mental*, incluso si la mente es idéntica al cerebro. Veamos a continuación al-

⁶ Aunque no exploraremos en este trabajo la naturaleza de estas explicaciones, podría tratarse de *explicaciones con causación internivel*. Un defensor de este tipo de explicaciones es Bermúdez (2003), quien partiendo del "argumento a partir de la conducta habilidosa" (*skilled behavior*) defiende la existencia de casos que requieren explicaciones que postulen causación de nivel personal y sub-personal.

gunas razones adicionales a favor del empleo del uso del vocabulario intencional para las explicaciones de los trastornos psiquiátricos, en particular, para el caso de los delirios.

Nivel personal: el “espacio de las razones”

Adoptar un vocabulario intencional para dar cuenta de los trastornos psiquiátricos puede brindar ciertas ventajas explicativas y predictivas (Bortolotti, 2005; Graham, 2010). En el caso de los delirios, esto es particularmente manifiesto, pues no se trata solo de dar cuenta de la etiología de una creencia irracional,⁷ sino además de explicar: (i) por qué la creencia delirante se mantiene, es decir, por qué el sujeto es insensible a la evidencia contraria que pueden proveer otros agentes; (ii) los fallos de racionalidad epistémica en relación con la incapacidad de algunos sujetos delirantes para *justificar* sus creencias delirantes, *i.e.*, para proveer razones a favor de las mismas; y (iii) los fallos en la racionalidad procedimental, pues los sujetos a veces no actúan en coherencia con sus delirios.

Si bien se podría explicar por qué la creencia se mantiene, apelando a algún tipo de sesgo cognitivo de nivel subpersonal,⁸ esto no bastaría para dilucidar por qué el sujeto no está *epistémicamente justificado* para sostener las creencias que tiene o por qué, en caso de esgrimir algunas razones a favor de tales creencias, no son consideradas “buenas razones” por el resto de la comunidad epistémica; es decir, por qué las creencias del sujeto no se rigen por los principios de justificación pública.

Para estos propósitos resulta relevante especificar el contenido semántico de la creencia delirante, rastrear las relaciones entre distintos estados mentales y evaluar los vínculos de estos estados con

⁷ Si bien actualmente no hay ninguna hipótesis sólida sobre cómo se originan las creencias delirantes (al menos no en el nivel de descripción neuronal), existen hipótesis como la del aumento del tono dopaminérgico. En los casos de trastornos psicóticos, se afirma que habría una hiperactividad dopaminérgica que generaría una *saliencia aberrante*. Según esta hipótesis, la liberación de dopamina en el cerebro puede llevar a que una persona dote de una relevancia e importancia absurda cualquier estímulo banal, como producto de una alteración dopaminérgica. Como algunos pacientes psiquiátricos manifiestan el delirio solo en situaciones extremas (por ejemplo, en brotes psicóticos), se postula que podría ser el tono hiperdopaminérgico lo que daría la saliencia aberrante a estas ideas (Kapur, 2003).

⁸ Algunos autores señalan que esto se explica por la presencia de sesgos cognitivos exacerbados. Por ejemplo, “*jumping to conclusions*”, sesgo de confirmación, de atribución externa, necesidad de cierre, etcétera (Colbert & Peters, 2002; Davies, Coltheart, Langdon & Breen, 2001).

las acciones del sujeto. Por ejemplo, en los casos de síndrome de Capgras, cabe preguntarse: ¿por qué el sujeto termina creyendo que su cónyuge es un impostor y no un alien o un robot? Y en los casos de síndrome de Cotard: ¿por qué el sujeto cree que está muerto y no, por ejemplo, que es una estatua? Asimismo, en el caso de los delirios *motivados* esto resulta de especial importancia. Este tipo de delirios son aquellos que emergen por los beneficios psicológicos que confieren al sujeto, siendo respuestas psicológicas activas a conflictos altamente perturbadores. La creencia delirante protegería al sujeto de auto-representaciones negativas que podrían llevar a depresión, baja autoestima y emociones negativas (Bortolotti, 2015; Bell, 2003; Bentall, Corcoran, Howard, Blackwood & Kinderman, 2001; Pavlickova *et al.*, 2013). En estos casos, similares a los de auto-engaño,⁹ rastrear el contenido semántico de tales creencias y las posibles explicaciones motivacionales (realizadas en términos de relaciones entre creencias y deseos) resulta importante para comprender la emergencia y mantenimiento de la creencia delirante. Esto puede permitir, además, colegir el rol que está jugando el síntoma en el trastorno del sujeto e incluso en la terapia; por ejemplo, en terapias de tipo cognitivo-conductual (CBT) posibilitaría adoptar una actitud crítica hacia el contenido. Para hacer inteligibles todas estas cuestiones, es preciso recurrir a un vocabulario intencional que nos permita explicar -aunque con restricciones- los pensamientos y comportamientos del sujeto delirante, apelando a criterios normativos particulares (diferentes de los biológicos) que tienen que ver con la racionalidad del agente.

Por otra parte, también es necesario apelar a criterios no biológicos para dar cuenta de los delirios, por cuanto la definición de estos contiene una cláusula cultural. La definición provista por el DSM V, según vimos, afirma que “la creencia no debe ser ordinariamente aceptada por otros miembros de la cultura o subcultura de la persona (esto es, no es un principio de la fe religiosa)” (APA, 2013, p. 824). La definición requiere de propiedades extrínsecas. Esto tiene un sentido: admitir un enfoque sensible al contexto es algo que

⁹ Casos en los cuales el sujeto termina creyendo que p porque desea creer que p.

se ve apoyado por consideraciones pragmáticas, ya que previene la “patologización” de millones de individuos que tienen, por ejemplo, convicciones religiosas o ciertas supersticiones. Después de todo, este tipo de creencias parecen compartir muchas características con los delirios, en el sentido de que son firmemente sostenidas y resistentes a contra-evidencia, a la vez que son débilmente soportadas por evidencia. Entonces, desde un punto de vista pragmático, parece justificado añadir una cláusula que sirva para excluir del rango de posibles delirios las creencias extrañas dependientes de la cultura. Estas valoraciones, así como las de bienestar relativas a la disrupción en el funcionamiento cotidiano del sujeto (McKay, Langdon & Colheart, 2009), también dependen de otros criterios normativos que no son los de un buen funcionamiento cerebral.

En síntesis, podemos decir que retener cierto grado de simpatía hacia el naturalismo en las investigaciones filosóficas sobre los delirios es fructífero para desarrollar teorías sobre las creencias (ordinarias y delirantes) empíricamente informadas, que no resulten contradictorias con los desarrollos actuales en psicología cognitiva y neurociencias. Por ejemplo, sirve para abandonar nociones idealizadas de racionalidad que han prevalecido en la tradición filosófica, para comprender los procesos subpersonales involucrados en la formación y mantenimiento de creencias (ordinarias y “anormales”), para reconocer los sesgos cognitivos en la evaluación de la evidencia a favor y en contra de las creencias, para comprender nuestra capacidad de manejar las incoherencias o disonancias cognitivas, etcétera. Pero esta es solo una parte de la explicación, y las razones de nivel personal también resultan necesarias.

Como hemos señalado, los criterios normativos de racionalidad (en un nivel interno al sujeto), así como los criterios externos, sociales y pragmáticos (como los relativos a la incapacidad de proveer razones que se rijan por los criterios de justificación pública), no resultan inteligibles bajo una explicación puramente natural. Para hacerlos inteligibles, también debemos situar la mente en el “espacio de las razones” (McDowell, 2003). Cuando hacemos esto no damos una descripción del estado en sí, sino que, además, “lo estamos colocando en el espacio lógico de las razones, de justificar lo que uno

diga y ser capaces de justificarlo” (Sellars, 1971, p. 186). Cuando, por ejemplo, creemos que *p*, se espera que seamos capaces de dar razones a favor de nuestra afirmación. Este tipo de compromiso es inteligible en un esquema lógico (no uno meramente causal) que permita algún criterio de corrección. Es decir, ubicando el estado mental en un contexto *normativo*, “en el cual la creencia de que las cosas son de tal o cual modo, sea una actitud que puede ser *correcta* o *incorrecta* en función de si las cosas son de tal y cual modo” (McDowell, 2003, p. 16). Solo así podremos dar cuenta del contexto normativo bajo el cual juegan las creencias, algo que desde una perspectiva exclusivamente naturalista -tanto por parte de la filosofía como de la psiquiatría misma- no sería posible.

Conclusiones

A lo largo de este artículo hemos defendido que si se asume una perspectiva naturalista moderada para el estudio de los trastornos psiquiátricos, puede resultar fructífera la interacción entre filosofía y evidencia empírica. Esto puede permitir elaborar enfoques sólidos sobre los trastornos psiquiátricos, que tomen en consideración los desarrollos conceptuales de la filosofía de la mente, así como enriquecer la filosofía con datos clínicos sobre la vulnerabilidad de la mente humana.

De igual manera, señalamos que buscar una descripción natural de nuestras capacidades mentales puede ser compatible con aceptar explicaciones de carácter híbrido y no implica por sí mismo adoptar una posición reduccionista. Siguiendo la tesis de Graham, hemos intentado mostrar que existe un espacio teórico amplio para albergar compromisos fisicalistas sobre la ontología de lo mental sin identificarlos con trastornos cerebrales, como quienes adhieren a la concepción de la psiquiatría biológica.

Particularmente, en el caso de los delirios, argumentamos que para comprender las creencias delirantes *qua* delirantes, será preciso servirnos de un abordaje híbrido que pueda contemplar las causas naturales del fenómeno, así como su evaluación normativa. Evitando posturas radicales, como las de la psiquiatría biológica o las de

sus detractores (en particular, aquellos “normativistas” que rechazan de plano cualquier posición naturalista), podemos acercarnos a una postura híbrida que resulte más provechosa al postular explicaciones de causación internivel. Restará explorar el desarrollo del marco conceptual explicativo necesario para mostrar cómo podrían integrarse ambos tipos de explicaciones.

Por último, cabe mencionar las limitaciones de un enfoque tal. Como ha señalado Dennett (1998), sea cual sea el procedimiento que usemos para adscribir estados intencionales a una persona con el fin de explicar algún rasgo de la conducta, si la interpretación parece justamente no explicarlo, entonces debemos cambiar hacia un nivel más bajo de elucidación. Así, aunque el tipo de abordaje híbrido propuesto podría funcionar en muchos trastornos psiquiátricos, puede ser necesario abandonar el nivel personal en aquellos de gravedad (como algunos casos de esquizofrenia de tipo desorganizado), en los cuales se rompen seriamente los vínculos entre los estados mentales y, en consecuencia, los sujetos se alejan demasiado de los criterios de racionalidad. En estos casos, si la explicación intencional ya no es útil para racionalizar la conducta del sujeto, simplemente deberemos apelar a explicaciones de carácter subpersonal, haciendo referencia a causas puramente mecánicas/neuronales. 

Referencias

- American Psychiatric Association [APA] (2013). *Diagnostic and statistical manual of mental disorders: DSM-V-TM* (5th ed.). Washington, DC: American Psychiatric Association.
- Andreasen, N. (1984). *The biological revolution in Psychiatry*. New York: Harper & Row Publishers.
- Andreasen, N. (2001). *Brave new brain: Conquering mental illness in the era of the genome*. New York: Oxford University Press.
- Andreasen, N. (2005). *The creating brain*. New York: Dana Press.
- Arpaly, N. (2005). How it is not “just like diabetes”: mental disorders and the moral psychologist. *Philosophical Issues*, 1(15), 282-298.
- Balestri, M., Calati, R. & de Ronchi, D. (2014). Genetic modulation of personality traits: a systematic review of the literature. *International Clinical Psychopharmacology*, 1(29), 1-15.
- Bayne, T. & Pacherie, E. (2005). In defence of the doxastic conception of delusion. *Mind & Language*, 2(20), 163-188.
- Bell, D. (2003). *Paranoia*. Cambridge: Icon.
- Bentall, R., Corcoran, R., Howard, R., Blackwood, N. & Kinderman, P. (2001). Persecutory delusions: a review and theoretical integration. *Clinical Psychology Review*, 1(21), 1143-1192.
- Bermúdez, J. (2003). *Thinking without words*. Oxford: Oxford University Press.
- Blackwood, N., Howard, R., Bentall, R. & Murray, R. (2001). Cognitive neuropsychiatric models of persecutory delusions. *American Journal of Psychiatry*, 1(158), 527-539.
- Bortolotti, L. (2005). Delusions and the background of rationality. *Mind and Language*, 2(20), 189-208.
- Bortolotti, L. (2015). The epistemic innocence of motivated delusions. *Consciousness & Cognition*, 33(1), 490-499.
- Bortolotti, L. & Miyazono, K. (2015). Recent Work on the Nature and Development of delusions. *Philosophy Compass*, 10 (9), 636-645.
- Bracken, P., Thomas, P., Timimi, S. & Yeomans, D. (2012). Psychiatry beyond the current paradigm. *The British Journal of Psychiatry*, 6(201), 430-434.
- Charney, D., Nestler, E. & Bunney, B. (1999). *Neurobiology of mental illness*. Oxford: Oxford University Press.

- Chaufan, C. & Joseph, J. (2013). The “missing heritability” of common disorders: should health researchers care? *International Journal of Health Services*, 2(43), 281-303.
- Colbert, S. & Peters, E. (2002). Need for closure and jumping-to-conclusions in delusion-prone individuals. *Journal of Nervous and Mental Disease*, 1(190), 27-31.
- Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *The Quarterly Journal of Experimental Psychology*, 8(60), 1041-1062.
- Davies, M., Coltheart, M., Langdon, R. & Breen, N. (2001). Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry and Psychology*, 3(8), 133-158.
- Dennett, D. (1998). *La actitud intencional*. Barcelona: Gedisa.
- Graham, G. (2010). *The disordered mind: an introduction to philosophy of mind and mental illness*. London: Routledge.
- Graham, G. (2013). Ordering Disorder: mental disorder, brain disorder and therapeutic intervention. En K. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, G. Stanghellini & T. Thornton (Eds.), *Oxford Handbook of Philosophy of Psychiatry*. Oxford: Oxford University Press.
- Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 1(160), 13-23.
- Kapur, S., Phillips, A. & Insel, T. (2012). Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 12(17), 117-149.
- Kendler, K. (2012). The dappled nature of causes of psychiatric illness: replacing the organic functional/hardware software dichotomy with empirically based pluralism. *Molecular Psychiatry*, 4(17), 377-388.
- McDowell, J. (2003). *Mente y Mundo*. Salamanca: Editorial Sígueme.
- McKay, R. & Dennett, D. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, 6(32), 493-561.
- McKay, R., Langdon, R. & Coltheart, M. (2009). Sleights of mind: Delusions, and self deception. En: T. Bayne & J. Fernández (Eds.), *Delusion and Self deception: Affective and Motivational Influences on Belief Formation*. Hove: Psychology Press.
- Miller, J., Ding, S. & Sunkin, S. (2014). Transcriptional landscape of the prenatal human brain”. En: *Nature*, 1(508), 199-206.
- Moncrieff, J. & Middleton, H. (2015). Schizophrenia: a critical psychiatry perspective. *Current Opinion in Psychiatry*, 3(28), 264-268.

- Pavlickova, H., Varese, F., Smith, A., Myin-Germeys, I., Turnbull, O. & Bentall, R. (2013). The dynamics of mood and coping in bipolar disorder: longitudinal investigations of the inter-relationship between affect, self-esteem and response styles. *PLoS ONE Journal*, 4(8), 1-2.
- Read, J., Mosher, L. & Bentall, R. (2015). *Models of Madness Psychological, Social and Biological Approaches to Psychosis* (2nd ed.). London: Routledge.
- Sellars, W. (1971). *Ciencia, percepción y realidad*. Madrid: Tecnos.
- Stich, S. (1990). *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge: MIT Press.
- Stone, T. & Young, A. (1998). Delusions and brain injury. The philosophy and psychology of belief. *Mind and Language*, 1(12), 327-364.
- Szasz, T. (2008). *El mito de la enfermedad mental*. Buenos Aires: Amorrortu.
- Timimi, S. (2014). No more psychiatric labels: why formal psychiatric diagnostic systems should be abolished. *International Journal of Clinical and Health Psychology*, 3(13), 208-215.