

Delirios, intencionalidad y racionalidad: análisis crítico del enfoque de Donald Davidson

Emilia Vilatta

emiliavilatta@gmail.com

Licenciatura en Filosofía. Directora de TFL: Dra. Laura Danón
Recibido: 11/05/16 // Aceptado sin modificaciones: 20/06/16

Resumen

El presente trabajo tendrá por objetivo evaluar los alcances y límites del enfoque de Donald Davidson para ofrecer una interpretación intencional de sujetos que padecen distintos tipos de delirios. Como estrategia metodológica me propondré examinar este debate a la luz de las siguientes categorías diagnósticas: *trastorno delirante*, *esquizofrenia de tipo paranoide* y *esquizofrenia de tipo desorganizado*. Sostendré que la concepción davidsoniana puede extenderse sin mayores inconvenientes a sujetos diagnosticados con trastornos delirantes y esquizofrenia de tipo paranoide. Asimismo, siquiera parcialmente, también puede emplearse para dar cuenta de algunos casos de esquizofrenia de tipo desorganizado. No obstante, para que esta extensión sea posible, será preciso reconsiderar el holismo propuesto por Davidson y matizar algunas de sus afirmaciones respecto a las condiciones para la atribución de pensamiento.

Palabras clave: delirios, intencionalidad, racionalidad.

1. Introducción

Recientemente, el enfoque de Donald Davidson sobre la interpretación de la mente y la acción ha sido ampliamente criticado en trabajos filosóficos que apelan al fenómeno de los delirios psiquiátricos. Los críticos han argumentado que los sujetos que padecen de delirios no satisfacen los requisitos de racionalidad que, según Davidson, debería cumplir un agente para que sea legítimo atribuirle intencionalidad (Bortolotti 2005, Klee 2004). Sin embargo, arguyen, parece evidente que los sujetos delirantes tienen

estados intencionales con contenidos mentales, aunque éstos últimos puedan resultarnos bizarros y/o alejados de la realidad. Por lo tanto, concluyen que la concepción davidsoniana debe estar equivocada y que la atribución de pensamientos no requiere necesariamente la satisfacción de los requisitos especificados por Davidson.

El presente trabajo tendrá por objetivo examinar y evaluar los alcances y límites del enfoque de Davidson para ofrecer una interpretación intencional de los sujetos que



padecen distintos tipos de delirios presentes en diferentes trastornos psiquiátricos.

Como estrategia metodológica me propondré examinar este debate a la luz de algunas categorías diagnósticas, ya que considero que, a los fines de realizar un análisis adecuado, es preciso observar las particularidades de los distintos tipos de delirios en distintos cuadros psiquiátricos. En función de ello, tomaré en consideración las siguientes categorías diagnósticas: *trastorno delirante, esquizofrenia de tipo paranoide y esquizofrenia de tipo desorganizado*.

Defenderé la idea de que la concepción davidsoniana sobre los requisitos de racionalidad que debe satisfacer un agente intencional, puede extenderse sin mayores inconvenientes a sujetos con determinados delirios psiquiátricos, específicamente a aquellos que sufren de trastornos delirantes y esquizofrenia de tipo paranoide. Asimismo, siquiera parcialmente, también puede emplearse para dar cuenta de algunos casos de esquizofrenia de tipo desorganizado. No obstante, para que esta extensión sea posible, es preciso reconsiderar el holismo propuesto por Davidson y matizar algunas de sus afirmaciones respecto a las condiciones para la atribución de pensamiento.

En primer lugar, revisaré brevemente los requisitos que, según Davidson, debe satisfacer un sujeto para ser considerado un

agente intencional, racional y dotado de vida mental. Luego examinaré la crítica que ha formulado Lisa Bortolotti (2005, 2009) en contra del argumento del trasfondo defendido por Davidson en sus escritos sobre la irracionalidad. En contra de dicha crítica, argumentaré que el tipo de delirios examinados (delirios monotemáticos presentes en trastornos delirantes y, dentro de los mismos, especialmente el delirio de Capgras), no representan un problema para el enfoque davidsoniano ya que los sujetos que los padecen aún conservan un trasfondo de racionalidad.

Posteriormente en el cuarto apartado señalaré algunas distinciones diagnósticas que resultan de utilidad para el presente análisis e introduciré al debate la defensa davidsoniana esgrimida por Marga Reimer (2011, 2012). Allí argumentaré, en favor de Davidson, que incluso en los casos de esquizofrenia de tipo paranoide puede hallarse la suficiente coherencia y correspondencia como para considerar que satisfacen los estándares de racionalidad necesarios para la atribución intencional.

Por último, en el quinto apartado, me centraré en el caso de la esquizofrenia de tipo desorganizado. Siguiendo algunas consideraciones establecidas por Glock (2003) en torno al pensamiento animal, ofreceré una lectura alternativa del holismo

Davidsoniano que, según intentaré mostrar, permite seguir atribuyendo -al menos- algunas creencias simples, sin tener que dejar de tratar a estos sujetos como agentes dotados de algún grado de intencionalidad.

2. Sobre las condiciones de atribución de pensamientos en Donald Davidson

En la obra de Donald Davidson es posible hallar una serie de reflexiones que convergen hacia un proyecto central en el cual se reflexiona acerca de nuestra capacidad para interpretar los unos a los otros, aspecto profundamente vinculado a múltiples líneas de la investigación davidsoniana, tales como la comprensión de la naturaleza y del significado, de las actitudes proposicionales y de nuestra posición epistémica con respecto a nuestras mentes, las otras mentes y el mundo que nos rodea.

Reconstruir el ambicioso proyecto de Davidson sería no sólo arduo, sino tal vez imposible dados los límites evidentes de la presente investigación. Por lo tanto, mi objetivo será trazar esquemáticamente sólo algunos aspectos de su obra, que nos permitirán posteriormente discutir una posible explicación en clave davidsoniana de los delirios. En particular, me restringiré aquí a considerar los principios que propone Davidson como condiciones de posibilidad para todo acto de interpretación.

El principio fundamental para la interpretación, según Davidson, es el "*Principio de Caridad*". Este sostiene que, en su mayor parte, el hablante está en lo correcto, tanto sobre su entorno como en sus creencias en general y que es ampliamente racional (coherente). En el marco del proyecto de la interpretación radical, este principio es considerado no simplemente una suposición contingente, sino constitutiva (Ludwig, 2003).

En el proceso de interpretación, el intérprete se guiaría por tres supuestos:

(i) En primer lugar, habrá de asumir que los contenidos de las creencias más básicas del sujeto están constituidos por ciertos rasgos objetivos del entorno, los cuales causan dichas creencias en el sujeto.

(ii) En segundo lugar, y en relación con el primer supuesto, habrá de asumir que, en los casos más básicos, lo que el sujeto considera verdadero es lo que él mismo considera verdadero.

(iii) En tercer lugar, habrá de atribuir al sujeto la capacidad de pensar de modo coherente (de acuerdo con lo que el intérprete mismo considera como pensamiento coherente).

A menos que el intérprete acepte estos tres supuestos acerca del sujeto a interpretar, no

será capaz de dar sentido a las emisiones de este último.

Por otra parte, Davidson (1967, 1970, 1975, 1985, 1992, 2001, 2003, 2004) considera que sólo es posible atribuir un estado mental con un contenido específico en el marco de una densa red de otros estados mentales relacionados adecuadamente entre sí, ya que el contenido de cada estado mental particular depende de sus vínculos inferenciales con otros contenidos. O en otras palabras, la identidad de un pensamiento particular no puede separarse de su lugar en una red formada por otros pensamientos vinculados lógicamente entre sí.

Entonces, dado que para que la atribución de un estado mental determinado P sea posible es preciso que éste se encuentre adecuadamente integrado con otros estados mentales del agente (respetando la coherencia), resulta evidente que poseer una actitud proposicional particular implica tener un patrón de creencias que sea mayoritariamente coherente. Esta tesis holista se extiende del ámbito del pensamiento al de la acción intencional, la cual es entendida como "aquella acción que puede ser explicada en términos de creencias y deseos, cuyos contenidos proposicionales racionalizan la acción" (Davidson, 2003: 146)¹. Para que una acción sea intencional es

preciso que guarde vínculos de coherencia con una red más amplia de pensamientos y acciones del agente, según lo cual se espera que las acciones resulten inteligibles a la luz de sus creencias y deseos.

Para Davidson, la satisfacción del principio de caridad es condición de posibilidad de la interpretación, ya que no podemos atribuir inteligiblemente estados mentales a menos que estemos dispuestos a encontrar, en el patrón de conductas y actitudes, un amplio grado de racionalidad.

Siendo la racionalidad una restricción ineludible y no optativa en la atribución de estados intencionales a otro agente, cabe preguntarse entonces qué sucede con aquellos casos en los cuales identificamos pensamientos irracionales en un agente. ¿Es posible atribuir esos estados mentales irracionales y al mismo tiempo -según el principio de caridad- considerar que el agente es aún racional? Consciente de esta tensión, Davidson brinda algunas consideraciones respecto a por qué la tesis holista sobre el pensamiento sigue siendo válida aún en estos casos, y por qué la atribución puede realizarse aún sin mayores inconvenientes. Para ello, en sus escritos sobre la irracionalidad, (Davidson 1970, 1982, 1985, 2004) presenta un argumento, -denominado por Bortolotti (2005) "*argumento del trasfondo*"- destinado a

conciliar la existencia de fenómenos irracionales con su teoría de la interpretación. A su criterio, las creencias "nunca son irracionales en sí mismas, sino dentro de un patrón más amplio" (Davidson, 2004: 142). Al respecto, el argumento del trasfondo sostiene que sólo podemos interpretar a un agente como teniendo creencias irracionales, siempre que el mismo mantenga un trasfondo de racionalidad en el resto de sus estados mentales. Esto es así ya que sólo podremos dar sentido a tales creencias (e incluso a la designación de las mismas como "irracionales") por comparación con los demás estados mentales y acciones del sujeto, los cuales -según el principio de caridad- deben poder ser vistos como siguiendo las normas de racionalidad. No sólo por preservar la inteligibilidad, sino porque además, "sin el elemento de racionalidad, nos negamos a aceptar la explicación como apropiada para un fenómeno mental" (Davidson, 2004: 196).

3. Intencionalidad y racionalidad

Lisa Bortolotti (2002, 2003, 2004, 2005a, 2005b, 2009) se ha pronunciado en contra de la concepción davidsoniana de la atribución de estados intencionales. Su crítica resulta especialmente controvertida, dado que, en oposición a Davidson, niega la construcción

de racionalidad como necesaria para la atribución de estados intencionales. Con el objetivo general de desafiar la concepción davidsoniana, Bortolotti ataca el holismo de lo mental defendido por Davidson, centrándose en particular en el argumento del trasfondo.

Bortolotti sostiene que los sujetos afectados por delirios parecen ser genuinos creyentes y sus creencias pueden ser exitosamente explicadas en términos intencionales, pero que estos sujetos no satisfacen los criterios que, según Davidson, deben cumplir para que el trasfondo de racionalidad esté en su lugar. Su argumentación puede ser esquematizada del siguiente modo:

1. Si el argumento del trasfondo es correcto, entonces no podemos atribuir estados mentales irracionales a menos que el sujeto sea ampliamente racional (i.e. a menos que su patrón de estados mentales y acciones conserve un trasfondo de racionalidad).
2. Podemos atribuir estados mentales irracionales a sujetos que no son "ampliamente racionales" (i.e. que no conservan un trasfondo general de racionalidad). El caso de los delirios lo prueba.
3. Por lo tanto, la conclusión del argumento davidsoniano es falsa.

El desafío para Bortolotti es mostrar la verdad de la segunda premisa, es decir, que efectivamente podemos atribuir estados irracionales (ej: una creencia delirante) a sujetos que no son ampliamente racionales, es decir, que no conservarían un trasfondo general de racionalidad.

Para ello, Bortolotti examina un delirio monotemático²: el del síndrome de Capgras³ y establece algunas comparaciones respecto a casos de irracionalidad ordinaria (ej: una falacia). Su comparación pretende llevarnos a aceptar que en los casos de Capgras la falla de racionalidad es distinta -en un sentido específico- respecto a un fallo ordinario. Para ello, Bortolotti enfatiza un elemento particular de la racionalidad: la capacidad para revisar las propias creencias. Luego afirma que una de las condiciones que, de acuerdo a Davidson, se impone a los agentes intencionales, es que cuenten con la capacidad para *restaurar* la racionalidad, es decir para revisar sus creencias una vez que una desviación respecto a las normas de racionalidad ha sido señalada o reconocida. Al respecto, afirma que aunque se les puedan atribuir creencias, los sujetos delirantes no pueden considerarse racionales, en tanto sus creencias son siempre enérgicamente resistentes a la revisión. Y agrega: "el caso de los delirios proporciona una razón para cuestionar la idea aceptada de que no puede

haber intencionalidad sin racionalidad" (Bortolotti, 2009: 392).

Así, Bortolotti concluye que esta resistencia a la revisión de creencias por parte de los sujetos delirantes hace que deban ser vistos como carentes de un trasfondo general de racionalidad.

4. Una defensa de la concepción davidsoniana de la atribución intencional

Según vimos, Bortolotti asume que para Davidson, tener un trasfondo general de racionalidad implica ser capaz de llevar a cabo procesos adecuados de revisión de creencias. Pero, su argumentación parece asumir implícitamente que para ser racional es preciso satisfacer un requisito más fuerte que aquel que exige el propio Davidson. La argumentación de Bortolotti puede ser reconstruida del siguiente modo:

1. Ser un sujeto "ampliamente racional", es decir conservar un trasfondo de racionalidad en el patrón de estados mentales, implica ser capaz de revisar adecuadamente *todas* las creencias.
2. El sujeto que padece de síndrome de Capgras falla en la revisión de algunas de sus creencias: no corrige sus creencias delirantes.
3. Por lo tanto, el sujeto carece de un trasfondo general de racionalidad.

Sin embargo, antes de aceptar este argumento, cabe señalar dos puntos. Por un lado, la primera premisa parece demandar un requisito más fuerte del que exige Davidson, ya que, para éste, no tenemos problemas en comprender perturbaciones menores cuando hay un trasfondo de racionalidad que en gran medida compartimos, con lo cual el hecho de que alguien actúe en algunas ocasiones de manera irracional no pone en peligro la capacidad para considerarle un sujeto intencional ni para atribuirle estados mentales. Aunque un pensamiento o acción no siga los principios de racionalidad (como podría ocurrir en el caso de la creencia del sujeto delirante), en tanto la mayoría de sus otros pensamientos/acciones sí lo hagan, a criterio de Davidson, la inteligibilidad seguirá estando garantizada. Esto último se explica por la tesis de la "*compartimentación de lo mental*" que hay defendido Davidson en sus escritos sobre irracionalidad.

En la búsqueda de una explicación psicológica de cómo los creyentes pueden ser inconsistentes (en particular para el caso de la *akrasia*), Davidson introduce la idea de que la mente puede estar particionada. Al respecto sostiene que la mente puede estar dividida en dos o más estructuras cuasi-independientes. Al interior de cada una de estas partes, los eventos mentales se relacionarían de modos que respeten los

principios de la racionalidad. Sin embargo, un estado mental perteneciente a una parte, podría vincularse causalmente con un estado mental de otra parte, sin mantener vínculos racionales con él. La característica necesaria para que haya estructuras de este tipo, es que una parte de la mente exhiba un mayor grado de racionalidad o coherencia que la que se atribuye al todo. Es decir que, asumiendo que las partes tienen algún grado de independencia entre sí, es posible comprender cómo estas pueden albergar inconsistencias y, a la vez, interactuar causalmente unas con otras. Lo único que se requiere a tal fin es que cada parte este compuesta por elementos organizados, en cuyo interior haya un aceptable grado de coherencia, y que cada elemento pueda operar sobre otro según la modalidad de la causalidad no racional. De este modo, Davidson logra resguardar la coherencia como requisito del pensamiento, y a la vez dar cuenta de cómo el fenómeno de la irracionalidad es posible⁴.

Siguiendo estas ideas cabe pensar que la creencia delirante (ej: sobre el supuesto impostor) puede hallarse en alguna de las partes de la mente dividida y que, en un proceso global de revisión de creencias, este compartimiento mental resulta inmune a la evidencia proveniente de otros.

Lo único que necesitamos, para seguir considerando al agente como ampliamente racional (es decir, conservando un trasfondo de racionalidad), es que *la mayoría* de sus creencias, intenciones y acciones se ajusten a los principios de racionalidad. Ahora bien, pienso que esto se cumple en los casos de delirios de Capgras, es decir que el patrón de estados mentales y comportamientos del sujeto exhibe un trasfondo de racionalidad que se ajusta a las normas davidsonianas⁵. Este parece ser el caso ya que los sujetos incurren en inconsistencias explícitas, mantienen vínculos coherentes entre la mayoría de sus creencias y acciones y si bien las creencias delirantes son -por hipótesis- falsas, sigue siendo probable que la gran mayoría de sus otras creencias mundanas sean verdaderas. Así pues, mientras el resto de creencias, actitudes y acciones del sujeto se atengan a los principios de racionalidad, tenemos buenas razones para considerar que ese sujeto posee el trasfondo de racionalidad necesario para posibilitar la atribución intencional. En breve, si el delirio es circunscripto (como ocurre en el caso del Capgras), no comportará un problema para el enfoque davidsoniano, ya que éste acepta fallos de racionalidad de carácter local. Por otro lado, y de modo independiente, cabe examinar, en relación con la segunda premisa, si la creencia del sujeto delirante y la

supuesta evidencia disconfirmatoria efectivamente representan una violación explícita a las normas de racionalidad davidsonianas.

Si consideramos el carácter peculiar de la experiencia del sujeto que padece tal trastorno, veremos que, aunque a simple vista parece irracional, de hecho no incurre en una violación explícita de racionalidad.

Según ha sugerido Pacherie (2008), el sistema subyacente a los sentimientos de familiaridad que experimentamos ante rostros bien conocidos, es de carácter modular. En función de esto, el delirio presente en el síndrome de Capgras (su formación y persistencia) podría ser explicado por la presencia de fallos en este sistema modular emocional. La modularidad del sistema que produce la sensación de familiaridad (y no-familiaridad) podría contribuir a explicar porqué las creencias delirantes se mantienen firmemente.

Así, la falla de un proceso afectivo modular involucrado en el reconocimiento de los rostros (en particular: en las expresiones emocionales, la identificación de firmas dinámicas de los rostros y la generación de respuestas autonómicas como los sentimientos de familiaridad) puede permitir comprender las experiencias y creencias delirantes de los pacientes con Capgras. Si efectivamente en estos casos opera una falla

de este tipo, las creencias delirantes “heredarían su contenido de la experiencia delirante de falta de familiaridad” (Pacherie, 2008: 120).

Si este enfoque es correcto cabe concluir que la razón por la cual los pacientes con Capgras fallan para desestimar sus creencias delirantes no es que ellos no puedan utilizar estos procedimientos de comprobación de creencias, sino que éstos no logran producir evidencia disconfirmatoria. En otras palabras, el paciente Capgras no es epistémicamente incompetente; más bien es víctima de un círculo vicioso epistémico. Sin embargo, afortunadamente, este círculo vicioso se limita a las creencias con un tipo específico de contenido y etiología (la creencia delirante), de ahí el carácter circunscripto del delirio.

En síntesis, en los casos de delirio de Capgras los sujetos no fallarían en restaurar la racionalidad, sino que no encontrarían disconfirmatoria de peso como para verse obligados a revisar su creencia delirante. Esto se debe al carácter peculiar de la vivencia en primera persona que los sujetos experimentan. Pienso que, bajo este enfoque, tenemos buenas razones para desestimar la segunda premisa del argumento de Bortolotti.

4.1. *Precisando algunas distinciones diagnósticas*

En los debates recientes, los filósofos han tendido a tratar al delirio como un fenómeno homogéneo, sin tener en cuenta sus diversas manifestaciones en distintos cuadros psiquiátricos. Sin embargo, las características particulares de cada trastorno resultan de importancia para la discusión, por lo cual propondré aquí valernos de algunas distinciones presentes en el Manual Diagnóstico y Estadístico de los trastornos mentales, 4^{ta} edición (DSM IV, APA, 2000). Me restringiré a aquellos diagnósticos de trastornos mentales que han estado en discusión y en los cuales el síntoma del delirio aparece entre las características definitorias: trastorno delirante, esquizofrenia de tipo paranoide y esquizofrenia de tipo desorganizada. Veamos cuáles son sus características:

- *Trastorno Delirante*: se caracteriza por al menos 1 mes de ideas delirantes no extrañas, sin otros síntomas de la fase activa de la esquizofrenia. Sus interpretaciones de estos acontecimientos acostumbran a ser consistentes con el contenido de las creencias delirantes. Asimismo, una característica habitual de los sujetos con trastorno delirante es la aparente normalidad de su comportamiento y de su aspecto

cuando sus ideas delirantes no son cuestionadas o puestas en juego. Excepto por el impacto directo de las ideas delirantes o sus ramificaciones, la actividad psicosocial no está deteriorada de forma significativa y el comportamiento no es raro ni extraño (APA, 2000: 322).

- *Esquizofrenia*: es caracterizada como una alteración que persiste durante por lo menos 6 meses e incluye por lo menos 1 mes de síntomas de la fase activa (p. ej., dos [o más] de los siguientes: ideas delirantes, alucinaciones, lenguaje desorganizado, comportamiento desorganizado o catatónico y síntomas negativos). Los síntomas característicos de la esquizofrenia implican un abanico de disfunciones cognoscitivas y emocionales que pueden afectar la percepción, el pensamiento inferencial, el lenguaje y la comunicación, la organización comportamental, la afectividad, la fluidez y productividad del pensamiento y el habla, la capacidad hedónica, la voluntad y la motivación y la atención. Dentro de este diagnóstico, se pueden diferenciar subtipos: tipo paranoide, tipo desorganizado, tipo catatónico, tipo indiferenciado y tipo residual. Nos restringiremos aquí al tipo paranoide y al desorganizado.

- *Tipo paranoide*: su característica principal consiste en la presencia de ideas claramente delirantes o alucinaciones auditivas, en el contexto de una conservación relativa de la capacidad cognoscitiva y de la afectividad. Principalmente las ideas delirantes son de persecución, de grandeza o ambas, pero también pueden presentarse ideas delirantes con otra temática (ej: celos, religiosidad o somatización). Habitualmente, estos sujetos muestran un deterioro mínimo o nulo en las pruebas neuropsicologías y en los test cognoscitivos. En estos casos no hay lenguaje desorganizado, ni comportamiento catatónico o desorganizado, ni afectividad aplanada o inapropiada (APA, 2000: 292, 293).

- *Tipo desorganizado*: sus características principales son: el lenguaje desorganizado, el comportamiento desorganizado y una afectividad aplanada o inapropiada. El lenguaje desorganizado puede ir acompañado de tonterías y risas que no tienen una conexión clara con el contenido del discurso. La desorganización comportamental (por ej: falta de orientación hacia un objetivo) puede acarrear una grave disrupción de la capacidad para llevar a cabo actividades de la vida cotidiana. Si hay ideas delirantes y alucinaciones, son fragmentadas

y no están organizadas en torno a un tema coherente (APA, 2000: 293, 294).

Ahora bien, en cuanto al delirio de Capgras discutido previamente, este constituye un caso de *delirio monotemático*. Estos delirios han sido tipificados como síndromes de falsa identificación delirante y es posible incluirlos dentro de los "*trastornos delirantes*". Al respecto, señalamos que al ser estos delirios circunscriptos, el enfoque davidsoniano podría lidiar con los mismos, específicamente apelando a la tesis de la compartimentación de lo mental que, según vimos, permite aceptar fallos de carácter local.

Por otra parte, en lo que refiere a la esquizofrenia y en particular a los dos subtipos mencionados, Marga Reimer (2011, 2012), ha realizado algunas consideraciones en pos de una defensa del enfoque davidsoniano, que resultan de interés.

En cuanto al tipo paranoide, a partir del análisis de las características de algunos casos, la autora ha argumentado que los delirios presentes en los mismos no son tan intensos como se piensa⁶, ya que parece haber la suficiente correspondencia y coherencia en el patrón de estados mentales del sujeto y los delirios mantienen numerosos vínculos inferenciales adecuados con el trasfondo de los otros estados mentales del sujeto (Reimer, 2011). La autora argumenta que los sujetos diagnosticados

con este trastorno conservan aún un trasfondo de racionalidad, y que es por esto mismo que, en tanto intérpretes, podemos detectar y atribuirle sus respectivos delirios.

Asimismo, si tomamos en cuenta, como señalé previamente, la tesis de la compartimentación de lo mental, estos casos no suscitarían mayores inconvenientes siempre que las creencias irracionales del sujeto sean limitadas y circunscriptas a un área particular, es decir que no afecten por completo su vida mental. Como hemos afirmado, siguiendo a Davidson, aunque uno o algunos pensamiento o acciones no se ajusten a los principios de racionalidad (como podría ocurrir en el caso de las creencias paranoicas del sujeto delirante), en tanto la mayoría de sus otros pensamientos y acciones si lo hagan, la inteligibilidad estará garantizada. Con lo cual, aún en estos casos, la racionalidad sigue siendo un requisito para la atribución de contenidos mentales.

Ahora bien, ¿podemos decir lo mismo de los casos de esquizofrenia de tipo desorganizado? Este tipo de trastorno, presenta algunos desafíos al enfoque davidsoniano ya que las características peculiares del mismo hacen pensar que no lograrán satisfacer los requisitos estipulados por Davidson para considerarlos como agentes intencionales.

5. Esquizofrenia de tipo desorganizado: ¿seres sin mente (*mindless*)?

Aunque Davidson nunca ha mencionado explícitamente patologías como la esquizofrenia, las características de aquello que ha denominado "*flagrante irracionalidad*" (Davidson, 2004: 219), parecen tener algún parecido con este tipo de trastornos. Al respecto, Reimer ha sugerido que la satisfacción de los criterios para el diagnóstico de esquizofrenia de tipo desorganizado equivaldrían a la irracionalidad global davidsoniana. La autora señala que en este tipo de casos "no hay evidencia clara de conexiones lógicas ni de contenido mental (identificable), contenido del tipo que podría, al menos en principio, ser parte de un patrón coherente que corresponde a la realidad" (Reimer, 2011: 673).

Estos casos, a su juicio, resisten a la explicación davidsoniana mediada por la atribución de contenidos. Sin embargo, no considera que esto desafíe la concepción de Davidson, antes bien, considera que precisamente juega a favor de la misma. Veamos por qué.

Como ya he señalado, Davidson (1992) afirma que los errores o incoherencias deben ser *locales*, ya que el error masivo socavaría las bases mínimas requeridas para detectar errores e incoherencias locales al hacer

imposible la identificación de contenidos de las creencias e intenciones del sujeto.

Asimismo, afirma que cuando la inteligibilidad está tan disminuida por la inconsistencia percibida, "una acusación de inconsistencia perdería aplicación por la falta de contenidos identificables sobre los cuales ser inconsistentes" (Davidson, 2004: 219).

Al respecto, Reimer (2012) esgrime que, de acuerdo al enfoque davidsoniano, si hallásemos un caso donde la incoherencia y falta de correspondencia fueran generalizadas y patentes, cabría esperar que el intérprete ni siquiera pudiera identificar los contenidos mentales del sujeto a interpretar. Lo que parece seguirse de esto es que sería legítimo dudar de si el sujeto posee aún pensamientos, ya que desde una concepción interpretativista como la de Davidson, tener pensamientos equivale precisamente a ser interpretable por otro agente como teniendo pensamientos⁷.

En función de ello, Reimer sostiene que, ante la ausencia de un patrón lo suficientemente rico de estados mentales intencionales, tal como podría suceder en los casos de esquizofrenia de tipo desorganizado, el intérprete ordinario estaría precisamente en la posición predicha por el davidsoniano. Es decir, se encontraría "completamente desconcertado y no tendría ni idea acerca de lo que el paciente cree realmente (desea, o

intenta, etc.), con lo cual el desconcierto absoluto seguramente sería la única respuesta" (Reimer, 2012: 18).

Reimer considera que se puede negar que los pacientes que sufren esquizofrenia de tipo desorganizado tengan pensamientos sin que esto implique negar que aún podamos atribuir a tales pacientes estados mentales, pero sólo en algún sentido no literal, quizás metafórico. Posiblemente, esgrime la autora, un paciente in-interpretable con esquizofrenia de tipo desorganizado, tenga pensamientos "en el mismo sentido en que Davidson (1982) consideraba que los animales podrían tener pensamientos" (Reimer, 2012: 673). Luego afirma que tales explicaciones son consistentes con la hipótesis de que aquellos cuyas conductas pretendemos explicar son criaturas carentes de mente: "criaturas <mindless>, en el sentido de que no poseen pensamientos" (Reimer, 2012: 674).

Si bien resulta al menos llamativa la comparación que realiza Reimer entre los sujetos que padecen trastornos psiquiátricos y los animales, una pista para pensar el por qué de esta analogía, es que la autora podría estar percatándose de un parecido entre ambos casos, que tiene que ver con los problemas vinculados al mal uso del lenguaje (en el caso de los sujetos con trastornos

psiquiátricos) y a la falta del mismo (en el caso de los animales).

Es bien conocido el hecho de que Davidson establece una fuerte dependencia entre pensamiento y lenguaje. Esto lo lleva a considerar que no es posible tener pensamiento en ausencia de lenguaje y, como consecuencia de ello, a rechazar terminantemente la posibilidad de atribuir pensamiento a los animales. Por razones semejantes, rechaza también la atribución de pensamiento a niños pequeños que aún no dominan un lenguaje. En este marco, resulta plausible entender el razonamiento de Reimer: los sujetos con esquizofrenia de tipo desorganizado tampoco podrían tener pensamientos, ya que no serían usuarios *plenos* del lenguaje. Asimismo, es posible que Reimer considere que tanto en el caso de los animales, como en el de los sujetos que padecen esquizofrenia de tipo desorganizado, no es posible hallar un patrón conductual lo suficientemente rico y complejo que legitime la atribución intencional.

Si bien la estrategia de establecer una analogía entre el pensamiento animal y el pensamiento de los sujetos con esquizofrenia de tipo desorganizado, puede iluminar algunos problemas compartidos que dificultaran en ambos casos la interpretación, pienso que si se matizan

algunos principios del holismo davidsoniano, que parecen ser excesivamente exigentes, podemos arribar a conclusiones menos pesimistas sobre estos sujetos y aproximarnos a una concepción de la vida mental más adecuada que reconozca la limitación -bien documentada por muchas investigaciones en psicología cognitiva- de nuestras capacidades cognitivas.

En primer lugar, a diferencia del caso de los animales que no poseen lenguaje, en los casos de esquizofrenia de tipo desorganizado (incluso cuando hay escasa expresión lingüística) no es implausible pensar que los pacientes pudieran asentir o disentir ante las descripciones que como intérpretes pudiéramos hacer. Con lo cual, no deberíamos rechazar *a priori* la posibilidad de obtener por esta vía información sobre sus pensamientos ya que son sujetos humanos que alguna vez poseyeron lenguaje (en un sentido pleno). En este sentido, la atribución resultaría más legítima que la atribución intencional a animales carentes de lenguaje, al menos desde la perspectiva davidsoniana. En contraste, esto no sería posible de ningún modo en los casos de los animales y la sustitución de expresiones co-referenciales podría llevarnos, como señala Davidson, de atribuciones verdaderas a atribuciones absurdas o ininteligibles.

En segundo lugar, en relación al carácter holista de lo mental, si bien la afirmación de que los pensamientos son individuados por sus contenidos y, en consecuencia, por sus conexiones lógicas, resulta plausible e intuitiva, cabe preguntarse si esta tesis ha de aplicarse del mismo modo para todas las clases de pensamiento y sin ningún tipo de matiz o cualificación.

Para responder a esta cuestión, Glock (2003) ha examinado diversos principios holísticos que cabría distinguir analíticamente y ha señalado que estos son "o demasiado fuertes para el propio caso de Davidson, porque se oponen incluso a casos plausibles de pensamiento humano, o bien demasiado débiles, porque incluyen algunas formas de pensamiento animal" (Glock, 2003: 282), algo que el propio Davidson no quisiera aceptar.

El principio holístico más fuerte que se podría invocar es el que sigue:

$$(A) (aBp \ \& \ (p \Rightarrow q)) \Rightarrow aBq$$

Según este principio, un sujeto *a* cree que *p* y *p* implica que *q*, por lo tanto, el sujeto que cree que *p*, debería creer también que *q*. Es decir, debería creer *todas* las implicancias de sus creencias. Coincido con la apreciación de Glock acerca del carácter excesivamente restrictivo de este principio. El autor señala que, por ejemplo, puede haber humanos que pueden creer los axiomas de la geometría

Euclidiana sin estar creyendo todos los teoremas implicados por ésta.

Ahora bien, Davidson podría afirmar que para atribuir inteligiblemente un pensamiento a un agente *a*, éste no necesita realmente creer que *q*, sino sólo *ser capaz* de creer (o ser capaz de comprender) que *q*, es decir, que es posible conocer que *q*, pero no necesario. Así, se comprometería con un segundo principio, más débil:

$$(B) (aBp \& (p \Rightarrow q)) \Rightarrow \diamond aBq.$$

Ahora bien, Glock argumenta que incluso este principio puede ser desafiado: ¿por qué no debería haber gente que crea, por ejemplo, en los axiomas euclidianos sin ser siquiera *capaces de aprender* tales teoremas? De hecho, muchos humanos usualmente rechazan al menos algunas de las consecuencias lógicas de sus creencias, incluso cuando éstas les son señaladas.

Davidson no necesita suscribir a un principio como (A), ni siquiera a uno que esté modalmente cualificado, como (B), ya que él mismo ha afirmado que no hay una lista fija de cosas que alguien con un concepto determinado deba creer. En sus palabras: "no hay una lista fija de cosas que alguien con el concepto de árbol deba creer" (Davidson, 2003: 98). Esto parece significar también que "entre estas implicancias no hay una lista fija de cosas que el agente *a* deba ser capaz de

apreciar, y por lo tanto *a* necesita ser capaz sólo de apreciar *algunas* de las cosas implicadas por *p*" (Glock, 2003: 283). Esto podría expresarse del siguiente modo:

$$(C) aBp \Rightarrow \exists q((p \Rightarrow q) \& \diamond aBq)$$

De acuerdo con este principio, si un sujeto humano no es capaz de comprender *ninguno* de los teoremas implicados por los axiomas euclidianos, entonces sus creencias en los axiomas no tienen el mismo contenido que la creencia de un humano que sí sea capaz de entender (al menos algunos de) aquellos teoremas y, en esta medida, los dos tendrían diferentes creencias.

Lo más razonable, para el defensor del holismo, sería entonces aceptar un principio como (C). Ahora bien, si aceptamos, siguiendo la argumentación de Glock, que (C) es el principio holista que debe satisfacer una criatura para que podamos atribuirle legítimamente estados mentales intencionales, entonces podemos mostrar que, de modo análogo a lo que ocurre con ciertos animales no humanos, también los sujetos diagnosticados con esquizofrenia de tipo desorganizado podrían cumplir con este requerimiento.

Lo que pretendo señalar es que es bastante probable que en cualquiera de estos casos pudiésemos encontrar al menos algunos momentos en los que los sujetos efectúen

acciones que resulten mínimamente coherentes, las cuales podamos explicar a la luz de sus creencias y deseos. Por ejemplo, cuando los sujetos manifiestan deseos de comer y luego realizan las acciones necesarias para ello (ej: sentarse, cortar la comida, ingerirla, etc.), o cuando manifiestan la necesidad de ir al baño y se dirigen hacia el mismo o piden que se los ayude; sólo por nombrar algunos casos de patrones de acciones cotidianas y básicas que posiblemente los sujetos realicen y respecto a los cuales le podamos atribuir estados intencionales (deseos, percepciones y/o creencias básicas) como causantes de las mismas. Lo que procuro enfatizar es que estas acciones son posibles y que, en este sentido, la atribución de pequeñas redes de pensamiento es una opción abierta que no deberíamos descartar *a priori* y que eventualmente podría ser legitimada apelando a evidencia empírica sobre las capacidades de los sujetos.

Entonces, si (C) es un principio holístico plausible, podemos decir junto a Glock, que existen redes de pensamientos más amplias y redes más pequeñas. Qué tipo de red se requiera en cada caso es algo que dependerá de la creencia y de la criatura en cuestión. En esta misma dirección, pienso que podemos aceptar que tales redes podrían presentar una coherencia y

correspondencia limitadas (menos plena que las redes de sujetos humanos adultos "normales"), en consonancia con su extensión y complejidad.

En consecuencia, del hecho de que los animales y algunas criaturas (pre-lingüísticas, o no plenamente lingüísticas) carezcan de redes tan ricas y complejas como las nuestras, no se sigue que no tengan creencias de ningún tipo. Por lo tanto, los sujetos diagnosticados con esquizofrenia de tipo desorganizado podrían presentar redes más pequeñas de estados mentales vinculados inferencialmente entre sí y podrían poseer, en consecuencia, pensamientos menos sofisticados que los de aquellos humanos adultos lingüísticamente competentes (sin trastornos) que poseen redes sumamente complejas y sofisticadas.

6. Conclusiones

A lo largo de este trabajo me he ocupado de examinar algunos debates recientes que se sitúan en la intersección entre la Filosofía de la Mente y la Psiquiatría. En particular, me he concentrado en el fenómeno de los delirios como contrapunto para discutir algunas tesis centrales del pensamiento de Donald Davidson.

El objetivo de la presente investigación fue evaluar los alcances y límites del enfoque de

Davidson para ofrecer una interpretación intencional de aquellos sujetos que padecen diversos tipos de delirios presentes en diferentes trastornos psiquiátricos.

Como estrategia metodológica me propuse examinar las discusiones en cuestión a la luz de distintas categorías diagnósticas.

He defendido la idea de que el enfoque davidsoniano puede lidiar con los casos de trastornos delirantes (con delirios monotemáticos como el síndrome de Capgras), con casos de esquizofrenia de tipo paranoide en los cuales se presentan delirios politemáticos mientras resulten circunscriptos, e incluso -al menos con algunos- casos de esquizofrenia de tipo desorganizado.

Para ello, en primer lugar, desestimé la crítica de Bortolotti al argumento del trasfondo davidsoniano. Allí argumenté que un examen cuidadoso de los delirios monotemáticos, en particular del delirio de Capgras, nos da razones para pensar que no se trata de que el sujeto falle en restaurar la racionalidad sino que, debido al mismo trastorno, no encuentra evidencia disconfirmatoria de peso como para verse obligado a revisar su creencia.

Luego brindé algunas consideraciones respecto a por qué puede decirse que aún en este tipo de casos el trasfondo de racionalidad se encuentra conservado.

Concluí así que las condiciones mínimas para que la atribución de estados intencionales tenga lugar en estos casos, se encuentran garantizadas.

En cuanto a los casos de esquizofrenia de tipo paranoide, sostuve, siguiendo a Reimer, que mientras los delirios paranoides sean circunscriptos y no afecten totalmente la vida mental del sujeto, podrá hallarse en los sujetos que los padecen la suficiente coherencia y correspondencia como para considerar que cumplen los estándares de racionalidad necesarios para atribuirles contenidos mentales.

Finalmente, concluí que la presuposición de racionalidad sigue siendo una restricción necesaria para la identificación y atribución de estados mentales intencionales.

Por último, me referí a los casos de esquizofrenia de tipo desorganizado, trastorno que fue pensando por algunos críticos como particularmente desafiante para el enfoque davidsoniano.

Intenté mostrar que no tenemos buenas razones para negar *a priori* la posibilidad de atribuir legítimamente al menos algunos estados intencionales a estos sujetos, dado que podrían conservar pequeñas redes de estados mentales, manteniendo mínimamente los requisitos de racionalidad. Aquí me propuse defender, apoyándome en las consideraciones de Glock, que tenemos

buenas razones para asumir un holismo que no sea tan radical como aquel con el que se compromete Davidson.

Si bien la tesis del holismo de lo mental parece correcta, así como también la postulación de los principios de racionalidad, ni el holismo ni la satisfacción de los principios de racionalidad necesitan ser *plenos*, algo que el mismo Davidson admite al plantear la tesis de la compartimentalización de lo mental. Con lo cual, el enfoque no tiene problemas para lidiar con los fallos de racionalidad que pueden presentar los sujetos delirantes, en la medida en que estos mantengan un trasfondo de racionalidad mínimo que garantice la inteligibilidad.

A pesar de esto, no podemos extraer sin más la conclusión de que el enfoque davidsoniano puede lidiar con *todos* los casos de sujetos delirantes. Las conclusiones a las que he arribado, antes bien, sólo buscan apuntar a una gradualidad en la posibilidad de extender el enfoque, pues las dificultades con las que se enfrenta la concepción de Davidson frente a los delirios psiquiátricos difieren según la gravedad de cada caso.

He intentado mostrar que no hay argumentos *a priori* para negar la atribución incluso a seres con esquizofrenia de tipo desorganizada. No obstante, la cuestión de *hasta dónde* podría extenderse

razonablemente el enfoque de Davidson (incluso bajo la versión revisada del holismo que he propuesto) no puede resolverse de antemano. En este sentido, será precisa la indagación empírica, y el análisis de registros clínicos para establecer qué tipos de dificultades hallamos en los diversos casos que se alejan del estándar de un humano adulto sin trastornos. No obstante, estas cuestiones tampoco son totalmente empíricas, ya que subsiste aún la cuestión filosófica respecto a qué evidencias considerar como válidas y/o relevantes para decidir si una criatura posee estados mentales.

7. Agradecimientos

El presente trabajo sintetiza los puntos centrales de mi trabajo final de licenciatura en Filosofía. Quisiera agradecer a Laura Danón por la dirección en dicho proyecto, también a Carolina Scotto, Aarón Saal y Mariela Aguilera por sus valiosos comentarios en calidad de tribunal examinador.

8. Notas

1. Cabe destacar que Davidson (1980) en "*Acciones, razones y causas*" defiende que las razones a la luz de las cuales actuamos son *causas* de las acciones, y que la explicación de la acción es una especie de explicación *causal*. Tales explicaciones de la acción citan pares "creencia-deseo" que causan la acción de modo conjunto y operan como razones a favor de la misma, desde el punto de vista del agente. Por ejemplo, un deseo especifica un fin que el agente tiene, y la creencia vincula alguna acción en particular a alguna posibilidad de alcanzar el fin. Davidson denomina a tales explicaciones "*racionalizaciones*".

2. Los delirios monotemáticos son aquellos que se hallan circunscriptos a un tema particular. En contraposición, los delirios politemáticos son aquellos pueden referir a más de una temática.

3. El síndrome de Capgras fue descrito por vez primera en 1923 por Capgras y Reboul-Lachaux bajo la denominación de "l'illusion des sosies" ("ilusión de dobles"), como un cuadro de falsa identificación delirante que consiste en la creencia de que una persona, generalmente cercana y afectivamente significativa, ha sido sustituida por un doble,

quien es considerado un impostor. Implica la convicción de que el familiar, aunque idéntico al real, es una persona diferente en algo (Madoz Gúrpide & Hillers, 2010).

4. Esto abre la posibilidad de que podamos atribuir a un agente conjuntos de creencias inconsistentes entre sí, pero consistentes al interior de cada uno de ellos (Davidson, 2004: p.178).

5. Es decir, que si bien puede haber un compartimento con creencias delirantes (sobre el supuesto impostor) que pudiera resultar inconsistente con el conjunto mayor de creencias del sujeto, ambos conservan un trasfondo de racionalidad que hace que sea posible ver al conjunto global de estados mentales del sujeto como respetando aún los principios de racionalidad.

6. Un autor que ha sugerido que los delirios de este tipo de trastornos son tan "intensos" que el enfoque davidsoniano no podría explicarlos es Klee (2004).

7. Al respecto afirma: "si otras personas o criaturas se encuentran en estados que no puedan describirse por estos métodos, no puede ser porque nos fallen nuestros métodos, sino porque esos estados no deberían llamarse estados mentales" (Davidson, 2003: p. 73).

9. Bibliografía

American Psychiatric Association & American Psychiatric Association (2000) *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (4th ed., text revision.), Washington, DC: American Psychiatric Association.

BORTOLOTTI, Lisa (2002) "Marks of Irrationality" en Lyons & Clarke (eds.), *Recent Themes in the Philosophy of Science*, Berlín: Kluwer Academic Publishers.

BORTOLOTTI, Lisa (2003) "Inconsistency and interpretation", *Philosophical Explorations*, vol. 6, Año 2, 109-123.

BORTOLOTTI, Lisa (2004) "Can We Interpret Irrational Behavior?", *Behavior and Philosophy*, vol. 32, Año 2, 359 - 375.

BORTOLOTTI, Lisa (2005a) "Intentionality without rationality", *Proceedings of the Aristotelian Society*, vol 105, Año 1, 369-376.

BORTOLOTTI, Lisa (2005b) "Delusions and the Background of Rationality", *Mind and Language*, vol. 20, Año 2, 189-208.

BORTOLOTTI, Lisa (2009) *Delusions and other irrational beliefs*, New York: Oxford University Press.

DAVIDSON, Donald (1967) "Truth and Meaning", *Synthese*, vol 17, pp. 304-23.

DAVIDSON, Donald (1970) "How is Weakness of the Will Possible?" en: Feinberg, Joel (ed.), *Moral Concepts*, Oxford: Oxford University Press.

DAVIDSON, Donald (1975) "Thought and Talk" en Guttenplan, Samuel (ed.), *Mind and Language*, Oxford: Oxford University Press.

DAVIDSON, Donald (1980) *Essays on Actions and Events*, New York: Oxford University Press.

DAVIDSON, Donald (1982) "Two Paradoxes of Irrationality" en Wollheim, Richard & Hopkins, James (eds.), *Philosophical Essays on Freud*, Cambridge: Cambridge University Press.

DAVIDSON, Donald (1985) "Incoherence and irrationality", *Dialectica*, vol 39, pp. 345-354.

DAVIDSON, Donald (1992) *Mente, mundo y acción: claves para una interpretación*, Barcelona: Paidós.

DAVIDSON, Donald (2001) *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.

DAVIDSON, Donald (2003) *Subjetivo, intersubjetivo, objetivo*, Madrid: Cátedra.

DAVIDSON, Donald (2004) *Problems of rationality*, New York: Clarendon Press.

GLOCK, Hans (2003) *Quine and Davidson on Language, Thought and Reality*, Cambridge: Cambridge University Press.

KLEE, Robert (2004) "Why Some Delusions Are Necessarily Inexplicable Beliefs", *Philosophy, Psychiatry, & Psychology*, vol 11, Año 1, 25-34.

LUDWIG, Kirk (2003) *Donald Davidson*. Cambridge, UK: Cambridge University Press.

Madoz Gúrpide, Agustín y HILLERS-RODRIGUEZ, Rosalía (2010) "Delirio de Capgras: una revisión de las teorías etiológicas", *Revista Neurología*, vol 50, Año 7, 420-430.

PACHERIE, Elizabeth (2008) "Perception, emotions and delusions: Revisiting the Capgras Delusion" en T. Bayne & J. Fernandez, *Delusions and Self-Deception*, UK: Psychology Press.

REIMER, Marga (2011) "A Davidsonian perspective on psychiatric delusions", *Philosophical Psychology*, vol 24, Año 5, 659-677.

REIMER, Marga (2012) "Davidsonian holism in recent philosophy of psychiatry" en Gerhard Preyer (ed.), *Donald Davidson on Truth, Meaning, and the Mental*, Oxford: Oxford University Press.