

teorema

Vol. XXVI/3, 2007, pp. 145-159

ISSN: 0210-1602

Autoengaño y responsabilidad

Jesús Antonio Coll Mármol

ABSTRACT

In recent years there has been an important debate between intentionalists and non-intentionalists on the nature of self-deception. Though both positions offer different accounts of this phenomenon, they share the idea that we are responsible of self-deception. However, both positions have their own problems to explain appropriately not only the kind of responsibility we attribute in cases of self-deception, but also whether we are responsible at all. In this article I will defend that we can capture the right kind of responsibility from a non intentional perspective. However I will argue that, contrary to some non-intentionalist accounts —like Mele’s— that the attribution of responsibility to self-deceivers presupposes their previous knowledge of the truth

RESUMEN

En fechas recientes ha habido un importante debate acerca de la naturaleza del autoengaño entre intencionalistas y no intencionalistas. Aunque ambas posiciones ofrecen diferentes explicaciones del fenómeno, comparten la idea de que somos responsables del autoengaño. Sin embargo, ambas posiciones tienen problemas para explicar apropiadamente, no sólo el tipo de responsabilidad que atribuimos en casos de autoengaño, sino también para dar cuenta de si somos siquiera responsables de dicho fenómeno. En este artículo defenderé que podemos capturar el tipo correcto de responsabilidad desde una perspectiva no intencional. Sin embargo, argüiré, contrariamente algunas posiciones no intencionalistas —como la de Mele— que la atribución de responsabilidad a quien se engaña presupone el conocimiento previo de la verdad por parte de éste.

El autoengaño es un fenómeno ampliamente extendido. Una madre deja de percibir que su hijo consume drogas a pesar de una abundante evidencia e información a favor. Una mujer sigue considerando fiel a su marido a pesar de que las evidencias apuntan a que las noches que supuestamente se queda trabajando en la oficina las pasa en un misterioso piso del centro de la ciudad. Un muchacho piensa que la chica que le gusta se está haciendo la dura y la vuelve a invitar a salir con él, a pesar de que es evidente para todos que ella pasa olímpicamente de él. Ante todos estos casos hay una respuesta clara que nos explica lo sucedido: los protagonistas son víctimas de un autoengaño. El

concepto de autoengaño tiene, de este modo, un carácter esencialmente explicativo. En concreto nos explica cómo es posible que alguien pueda haber llegado a creer una falsedad ante una evidencia en muchos casos apabullante en contra¹. Pero, ¿qué es exactamente lo que explica el concepto? ¿Qué sucede en un caso típico de autoengaño?

Una primera aproximación al contenido del autoengaño consiste en concebirlo, tomando literalmente el significado de la propia palabra, como un caso de engaño peculiar en el que engañador y engañado se funden en una misma persona. Así, el engaño interpersonal serviría como modelo para comprender el autoengaño. El problema es que el paralelismo con el caso interpersonal es problemático y enturbia en algunos casos más que aclara el caso del autoengaño. El caso de engaño interpersonal involucra al menos a dos personas, una que ejerce de engañador (A) y otra de engañado (B). Parece que, al menos en el caso más paradigmático, el engaño interpersonal involucra los siguientes elementos:

- a) A sabe que p.
- b) A intenta hacer creer a B que no-p.
- c) B termina creyendo que no-p a causa de A.

El problema de trasladar tal cual el caso interpersonal al personal es que parece convertirlo en un imposible. En el caso interpersonal podemos hablar de que engañador y engañado tienen creencias contradictorias una vez culminado el engaño, pero en el caso personal no es tan sencillo atribuir dichas creencias a la misma persona sin hacerle creer una contradicción flagrante, algo que desafía nuestros patrones de atribución de racionalidad. Además, encontramos el problema de la intencionalidad del engaño en el caso interpersonal. En dicho caso parece no sólo habitual sino absolutamente necesario atribuir a A la intención de hacer creer una falsedad a B para que tenga sentido hablar de engaño, pues parece que en caso de que no exista dicha intención no consideramos que se nos esté engañando, sólo que se nos ha inducido a error de un modo involuntario, a lo sumo negligente. Por otro lado, esa intención debe ser desconocida de modo crucial por B, pues el conocimiento de dicha intención le pondría sobre aviso de lo que A estaba tramando, convirtiendo en imposible la tarea de engañarlo por parte de A. Nada de esto parece ser posible en el caso personal, pues se requiere que el sujeto del autoengaño tenga por un lado la intención de engañarse y, al mismo tiempo, desconozca que tiene dicha intención pues, de lo contrario, el autoengaño sería imposible.

El problema de definir el autoengaño desde el paralelismo con el caso interpersonal ha generado dos posiciones en torno al problema de su definición: aquellas que mantienen, a pesar de los problemas, el paralelismo entre

ambos casos —intencionalistas— y las que niegan dicho paralelismo y mantienen que el caso de autoengaño no requiere de alguno de los elementos característicos del caso interpersonal, en especial de la intencionalidad —no intencionalistas—. Un asunto importante es que, a pesar de las divergencias en la explicación del caso, para todos ellos resulta claro que quien se engaña es, de un modo u otro, responsable de su estado, queriendo significar con esto que podría haberlo evitado si se hubiera tomado la molestia de hacerlo. Esto ha de ser interpretado como una consecuencia deseable de cualquier aproximación al autoengaño pues, si no se fuera responsable de dicho estado, si no se tuviera control sobre el mismo, la racionalidad de la creencia se vería seriamente afectada pues, ante cualquier creencia que uno pudiera tener, nunca se estaría seguro de si se trataba de una creencia que perseguía la senda de la verdad o más bien se trataba de un producto de procesos psíquicos autónomos sobre los que el sujeto no tiene control. Nos enfrentamos así ante una amenaza escéptica de primera magnitud, ante unos procesos que distorsionarían la verdad sin necesidad de genio maligno alguno y que se encontrarían dentro de nosotros mismos y contra los que sería imposible luchar.

No obstante, dicha apelación al control y responsabilidad por parte de ambas posiciones parece no estar totalmente justificada. A primera vista parece algo natural para los intencionalistas afirmar la responsabilidad del agente del autoengaño ya que si alguien se forma la intención de engañarse, está desde luego en su mano evitar el engaño. El problema que tienen los defensores del intencionalismo es, como veremos, que muchos de ellos, para explicar cómo este fenómeno es posible a pesar de que exista la intención de llevarlo a cabo, hacen que la intención de engañar permanezca oculta al agente, bien apelando a una partición de la mente o a diferentes voes o subagentes que se ocultan entre ellos información e intenciones. De este modo lo que parecería que en principio les daría derecho a declarar al agente del autoengaño responsable de su estado, finalmente les impide atribuir control alguno sobre este proceso a dicho agente.

Del mismo modo, parece que dicha ausencia de control que quien se engaña posee sobre su estado se sigue de las posiciones no intencionalistas pues, si se ha llegado a ese estado sin tener la intención de hacerlo, podría interpretarse que el autoengaño es algo que más bien se padece más que se elige. Y, si éste es el caso, quizá habría que compadecer a quien se engaña en lugar de atribuirle responsabilidad alguna sobre su estado. Esta es la posición de Neil Levy quien, en un reciente artículo [Levy (2004)], ha mantenido que desde posiciones no intencionalistas se sigue, al menos en casos paradigmáticos de autoengaño, una ausencia de responsabilidad por parte de quien se engaña, manteniendo a su vez que se trata simplemente “un tipo de error, y no presumimos que los que se equivocan sean culpables” [Levy (2004), p. 310]. Levy considera esto como una virtud de la posición no intencionalista, pero también podría verse como un defecto, más aún cuando los defensores de di-

cha posición insisten en mantener algún tipo de responsabilidad al respecto. De este modo, y al contrario de lo que aparentemente sucede con las posiciones intencionales, los no intencionalistas habrían de justificar y explicar qué grado de responsabilidad hay en el autoengaño según su aproximación al mismo.

En este artículo defenderé que desde posiciones no intencionalistas podemos capturar el tipo de responsabilidad que comúnmente atribuimos en casos de autoengaño y que no existe problema alguno en mantener simultáneamente que el autoengaño no es intencional y que está en nuestras manos evitar dicho estado. Defenderé que el problema reside más bien en mantener el carácter no intencional del autoengaño *junto* a la idea de que no es necesario para llegar a dicho estado un cierto conocimiento de que la evidencia con la que contamos apunta en la dirección contraria a lo que finalmente terminamos creyendo. Si se considera que éste es un requisito indispensable para que se dé el autoengaño, entonces podremos atribuir responsabilidad a quien se engaña en el sentido de ser negligente con respecto al escrutinio de la creencia a la que finalmente ha llegado, debido a que conocía información relevante en contra de dicha creencia.

I. INTENCIONALISMO Y RESPONSABILIDAD

Antes de pasar a abordar el problema de la responsabilidad en la posición no intencional, desearía examinar brevemente los problemas propios de la posición intencionalista. Tal como se ha dicho, es tentador analizar el fenómeno del autoengaño desde el modelo del engaño interpersonal. A ello parece invitar un análisis semántico del término. Al igual que en el caso interpersonal el engañador sabe que algo es falso y, sin embargo, hace creer a su interlocutor que cree que es verdadero con la intención de engañarle, en el autoengaño uno decide engañarse sobre un determinado asunto del que conoce o como mínimo fuertemente sospecha la verdad, intentando provocarse la creencia contraria a lo que sabe que es el caso. Los defensores de este modo de aproximarse al fenómeno defienden entonces como dato crucial que el autoengaño involucra la intención de causar en sí mismo una creencia falsa por parte de quien se engaña, así como un conocimiento tácito de la verdad del caso².

Hay que decir que no es sólo la definición interpersonal de autoengaño la que invita a considerar a este fenómeno como intencional. Aparte del relato de los ejemplos que se proponen, parece haber un dato que invita a abrazar dicha posición: el reconocimiento, no sólo por parte de los demás, sino también de nosotros mismos, que, de algún modo, tuvimos acceso a la verdad y la dejamos pasar en favor de algo que, desde el principio, teníamos razones para considerar falso³. Esto viene reforzado además por el hecho de que, en muchos casos, no vamos solos al fenómeno del autoengaño, sino que llega-

mos a ese estado después de tratar con los demás, dando razones de por qué ellos se equivocan y nosotros no. Después de tales explicaciones resulta difícil negar que se conociera de algún modo la verdad y que, debido a motivos que no deberían haber influido en el juicio, se buscara otra verdad más propia a dichos motivos.

Esta posición parece recoger claramente la idea de que quien se engaña es responsable de ello, pues, sabiendo o sospechando sobre la base de la evidencia disponible que algo era de un determinado modo, intenta creer algo a pesar de que, en el mejor de los casos, carece de razones para hacerlo. El tipo de responsabilidad es además máximo, pues, conociendo la verdad, intenta creer una falsedad para, por lo general, su propio beneficio. Las consecuencias que se sigan de su autoengaño son entonces algo que pertenece al propio agente, pues puede ser acusado no sólo de no haber evitado entrar en cierto estado, sino de haber entrado en él voluntaria y conscientemente a través de su intención de hacerlo.

Sin embargo no todo es tan sencillo. A pesar de sus atractivos, la propuesta intencionalista presenta, como vimos, al menos dos problemas conceptuales importantes. El primero de ellos es que el paralelismo parece obligarnos a atribuir al agente del autoengaño creencias contradictorias acerca de un mismo tema. En el caso interpersonal este problema no aparece, pues quien engaña sabe perfectamente qué es verdadero y qué falso, mientras que en el caso del autoengaño el agente ha de realizar simultáneamente tareas de engañador y engañado y ha de tener, como agente engañador, la creencia verdadera que causa el autoengaño y, como víctima de sí mismo, la creencia causada en el autoengaño. El segundo problema reside en cómo la intención de engañarse no evita el éxito del fenómeno. En el caso interpersonal, que el engañador tenga la intención de engañar no supone obstáculo alguno, pues el engañado desconoce dicha intención: si la conociera el engañado podría oponer resistencia y rechazar así el reto del engañador. Pero en el caso personal en el que uno mismo es engañador y engañado, el conocimiento por parte del engañado de sus propias intenciones debería convertir en imposible la tarea del autoengaño.

Ante estos problemas se pueden aducir ejemplos de autoengaño que son isomórficos con el caso interpersonal con la ayuda de lo que Mark Johnston denomina *medios autónomos* [Johnston (1988), p. 77]. Así, uno podría engañarse a sí mismo —y a los demás— tomando una droga que le permita olvidar lo que ha hecho en la última hora y, con ello, el delito cometido, pudiendo así despertar posteriormente ajeno a ningún sentimiento de culpa. Esto puede suceder así, pero no se trata en modo alguno de un caso típico de autoengaño. La madre que se engaña sobre sí su hijo toma drogas o el marido que descarta la infidelidad de su mujer contra toda evidencia no utilizan tales medios en el proceso de su engaño. Para explicar cómo es posible que en estos casos se produzca el autoengaño, el defensor de la intencionalidad parece

condenado a mantener que la intención de engañarse ha de permanecer de algún modo *oculta* a quien se engaña. Así, se apela a diferentes voes o subagentes que, aunque cargados de una racionalidad intrínseca, permanecen ocultos unos a otros [Rorty (1988)], o bien se habla de una partición de la mente, como hace Davidson, en la que una parte queda oculta para la otra [Davidson (1985)].

El problema que tienen todas estas soluciones es que sacrifican la responsabilidad del agente a favor de la posibilidad del fenómeno, convirtiendo en inoperante la intención de engañarse, pues una intención oculta es una intención a la que no se puede apelar para explicar la responsabilidad del agente en el caso. Si a esto unimos el hecho de que es el yo manipulador, o la parte separada de la mente a la que no tengo acceso, la que realmente conoce la verdad acerca del asunto, ¿cómo puede atribuírseme responsabilidad alguna sobre el estado en el que me encuentro? ¿No soy más bien víctima que culpable de mi estado? ¿Cómo podría haber evitado algo que ni tenía —o al menos no sabía que tenía— la intención de hacer ni sabía que era falso? Y, si lo sabía, ¿cómo pude tener una intención de engañarme o de creer algo sin al mismo instante descubrirme como engañador y así evitar el estado al que finalmente llego? El dilema para el intencionalista parece simplemente inevitable: el intencionalismo corre el riesgo de salvar la posibilidad del autoengaño creando un problema con respecto al control del mismo, dando así una imagen de la mente enrevesada en la que uno es víctima literalmente de sí mismo o, quizá, habría que decir de otro.

II. AUTOENGAÑO SIN INTENCIÓN Y RESPONSABILIDAD

Una posible solución al aparente callejón sin salida al que nos lleva el intencionalismo es la de no modelar el autoengaño desde el caso del engaño interpersonal, renunciando a aquellos supuestos que generan los problemas en el caso personal. Si podemos explicar el fenómeno del autoengaño sin atribuir la creencia en una contradicción o sin atribuir una intencionalidad que socavaría la posibilidad del estado que pretende explicar, esta posición habrá de ser valorada positivamente, pues, quizá, ponga al descubierto la verdadera naturaleza del fenómeno que estamos estudiando y que el lenguaje con sus juegos de parecido parece ocultar.

Alfred Mele es, sin lugar a dudas, el campeón de esta posición, ya que no sólo renuncia a la intención de engañarse como elemento explicativo del autoengaño, sino que renuncia asimismo a atribuir conocimiento previo de la verdad a quien se engaña. Ya en un primer libro [Mele (1987)] defendió que el autoengaño no debía ser modelado de acuerdo con el caso interpersonal y que aquellos que se engañan raramente lo hacen intencionadamente [Mele (1987), p. 123]. Pero es en un reciente libro donde Mele mejor define su po-

sición [Mele (2001), basado en Mele (1997)]. Según Mele los casos comunes de autoengaño pueden ser explicados sin apelar ni a que el agente del autoengaño mantenga creencias contradictorias, ni a que tenga intención alguna de engañarse. Lo decisivo en el autoengaño es la motivación del agente, que es la que da lugar a una creencia falsa. Así, en casos típicos de autoengaño lo que sucede es que el deseo de que algo sea el caso predispone al agente a recopilar información o a interpretar distintos fenómenos de un modo sesgado, provocando con ello la adquisición de una creencia falsa. ¿Es necesario suponer que esa persona tiene creencias contradictorias? Mele considera que no lo es. En muchos casos de autoengaño lo que ocurre es, según Mele, que el deseo de que *p* provoca que se adquiera la creencia de que *p* y la no adquisición de no-*p*, sin necesidad de atribuir conocimiento previo de la verdad, ni atribuir a dicho conocimiento un papel causal, como vimos que sí le otorgaban los intencionalistas. La idea fundamental reside en que en el deseo, y en algunos casos de lo que Mele llama casos *retorcidos* (*twisted*) de autoengaño⁴, las emociones son suficientes para predisponer a alguien a adquirir una creencia sin necesidad de suponer, además, que es el conocimiento de la verdad del asunto el que provoca el autoengaño.

¿Qué decir del requisito de la intencionalidad del autoengaño? Mele considera que también es innecesario. En casos típicos de autoengaño es mucho más fácil explicar cómo se llega a ellos sin suponer que se tiene la intención de engañarse, o la intención de adquirir una creencia falsa, que suponiéndolo. Desde luego que en la conducta de quien se engaña podemos ver estrategias para su autoengaño, estrategias que sí pueden conllevar intencionalidad, pero Mele considera más factible que dichas estrategias funcionen sin apelar a intención alguna de llegar a creer una falsedad, sobre todo porque dichas intenciones, de ser conscientes, volverían inefectivas dichas estrategias y, en caso de ser inconscientes, dejarían de explicar lo que supuestamente explicaban⁵.

La propuesta de Mele ha de ser considerada como esencialmente deflacionista. La explicación del fenómeno se hace sin apelar a factores normalmente considerados importantes a la hora de explicarlo, no sólo desde presuposiciones intencionalistas, sino también desde otras posiciones no intencionalistas. Así, por ejemplo, Mele critica la propuesta de Annette Barnes —que a su vez se inspira en M. Johnston— de que para explicar el autoengaño sea necesario presuponer siquiera que el agente desee reducir el estado de ansiedad que le generaría la posible verdad sobre determinado asunto [Barnes (1997)]. Mele niega [Mele (2001), p. 55] que tal componente afectivo sea un ingrediente inevitable en todos los casos de autoengaño. Todo lo que se requiere en tales casos es que no se sepa que *p*, se desee que *p*, que dicho deseo lleve al agente a tratar la información de manera distorsionada y que adquiera finalmente la creencia de que *p*. Ningún otro componente afectivo ni cognoscitivo es para Mele condición necesaria para explicar su estado.

Son varios los problemas que esta propuesta puede generar. Así, Mele ha sido acusado de no diferenciar entre el pensamiento desiderativo (*wishful thinking*) y el autoengaño, acusación que Mele acepta no como un problema de su posición, sino como una virtud de la misma [Mele (1987), pp. 134-135]. Pero quizá los problemas más graves de una posición deflacionista como la de Mele sean sin lugar a dudas los referidos al control de dicho estado. Esta ausencia de control puede ser vista desde diferentes ángulos. Así, Bermúdez acusa a Mele de no dar cuenta del carácter selectivo del autoengaño [Bermúdez (2000), p. 317]. Este autor considera que la apelación a la motivación del agente no explica por qué el autoengaño ocurre en una ocasión particular. Hay muchas situaciones en las que, por mucho que deseemos algo, no nos engañamos. Así, y dando por supuesto que tenemos control sobre nuestro estado de autoengaño, hay que suponer que el agente tiene la intención de engañarse para poder explicar el carácter selectivo del autoengaño. En caso contrario parece que Mele no puede explicar por qué algunos de nuestros deseos causan el estado y otros muchos no lo hacen.

La respuesta de Mele a Bermúdez no supone tanto una defensa del control sobre el estado de autoengaño como un ataque a la idea de que la intención de engañarse tampoco explicaría su carácter selectivo [Mele (2001), pp. 65-66]. Así, muchas veces la intención de hacer o conseguir algo finalmente no fructifica. ¿Qué diferenciaría los casos en los que conseguimos lo que intentamos y aquellos en los que no? No la intención, desde luego. Si es posible tener la intención de engañarse y no conseguirlo, este elemento no explicaría tampoco el carácter selectivo que Bermúdez atribuye al caso del autoengaño.

Pero, ¿tiene realmente el autoengaño ese carácter selectivo? ¿Tenemos realmente control sobre dicho estado? Neil Levy ha defendido que, en los casos comunes de autoengaño, carecemos del control adecuado para podernos declarar responsables. Levy mantiene que la idea de que el agente del autoengaño es responsable de su estado es un resto de la posición intencionalista sobre el autoengaño basada en una concepción cartesiana de la mente. Así, si abandonamos los requisitos que acompañan a la concepción tradicional del autoengaño —la idea de sostener creencias contradictorias y la intención de provocar el engaño— se seguiría que, en los casos comunes del fenómeno, no hay responsabilidad por parte del agente, precisamente porque no hay control sobre dicho estado. Si el autoengaño fuera algo que uno provocara de algún modo, uno podría evitarlo; pero si es algo que simplemente nos pasa, lo único que ocurre es que cometemos un error. Y aunque algunas veces la gente es responsable de sus errores, no hay la misma presunción en todos los casos [Levy (2004), p. 295].

Levy argumenta en contra de que, desde posiciones no intencionalistas, se siga una responsabilidad claramente definida para quien se engaña. Una vez que se ha abandonado el modelo intencional de autoengaño, que atribuía tanto el conocimiento de la verdad como la intención de engañarse, es mucho

menos claro que el autoengaño sea algo que uno decida llevar a cabo, más bien que algo que le *pase* a uno mismo. Según Levy se me puede culpar de las creencias falsas que mantengo siempre que se puedan remontar a un acto mío de negligencia epistémica conocida [Levy (2004), p. 304]. Levy no niega que en algunos casos de autoengaño pueda suceder esto, pero no tiene por qué ser así en todos, sobre todo si es posible que haya autoengaño sin los dos requisitos formulados por los intencionalistas. El punto principal de Levy es que, para atribuir responsabilidad a alguien acerca de cómo ha adquirido o mantenido sus creencias, es necesario que se den dos requisitos: que sea sobre una materia importante para él y, de modo decisivo, que se dude de algún modo de ella [Levy (2004), p. 305]. Se podría pensar que uno debe examinar con especial cuidado todas las creencias sobre asuntos importantes independientemente de las dudas suscitadas sobre las mismas, pero esto no es así para Levy, más aún si su posición se interpreta como el deber de tomar en consideración *todos* los argumentos o datos posibles en relación con una cuestión. Sólo cuando tengo dudas con respecto al asunto que me ocupa tengo el deber de hacerlo⁶. Pero, en el caso de las posiciones deflacionistas como la de Mele, quien se engaña carece de elementos para tener duda alguna respecto al tema sobre el que lo hace. Como dice Levy “el autoengaño efectivo parece excluir la satisfacción de (2) [tener dudas acerca del asunto sobre el que se engaña]” [Levy (2004), p. 307]. Mele piensa que sí que cabe algún control sobre los procesos que distorsionan nuestras creencias. Así, dice que los mecanismos distorsionadores que están en funcionamiento en el autoengaño están de algún modo bajo nuestro control y que se nos puede educar de modo que detectemos dichos mecanismos [Mele (2001), p. 103]. Levy no niega que algunos de estos mecanismos puedan reconocerse fácilmente, pero no cree que todos hayan de serlo, ya que alguno de los mecanismos descubiertos por la psicología social están lejos de ser intuitivos y conocidos por la mayor parte de la gente. Parece entonces que atribuir responsabilidad en el caso de autoengaño a quien lo sufre está lejos de ser claro.

Hay que señalar que Levy no piensa que esto sea una crítica para los no intencionalistas, sino una consecuencia realista de tal posición. El autoengaño es simplemente un tipo de error y no presumimos que quien se equivoca sea culpable. El problema es que tal conclusión debería figurar en el *debe* de dicha posición. Si el autoengaño es algo que se produce, al menos en algunos casos, sin intención de que se produzca y no tengo ningún dato que me permita reconocer que me encuentro ante una creencia distorsionada, ¿cómo puedo saber que las creencias en las que ahora confío para actuar de un determinado modo no han sido causadas de un modo inapropiado? ¿Cómo puedo saber si estas creencias han sido producidas por mecanismos mentales cuya formación no reconozco y, por tanto, soy víctima de dichos mecanismos? Si Levy estuviera en lo cierto con respecto al caso del autoengaño, nos encontraríamos, como dije en la introducción, ante una duda tan poderosa al menos co-

mo la del genio maligno cartesiano aunque, en aquel caso, el engaño se producía por alguien ajeno a mí mismo, y ahora el enemigo está en casa y no hay forma de descubrir sus tretas. El ideal de autonomía del agente y su responsabilidad con respecto a sus creencias se van por la borda.

III. EN DEFENSA DE LA RESPONSABILIDAD EN EL AUTOENGAÑO

En mi opinión Levy tiene razón en mantener que desde una posición deflacionista como la de Mele se sigue una ausencia de control por parte de quien se engaña y, por ende, una ausencia de responsabilidad. Pero lo que se seguiría de esto no es tanto que el autoengaño sea un proceso sobre el que no tenemos control, sino más bien que Mele se equivoca al explicar el autoengaño. La ausencia de control y de responsabilidad no puede ser considerada como un hallazgo virtuoso de una cierta forma de entender un fenómeno, sino como una razón para rechazar dicha posición. En casos comunes de autoengaño pensamos de los demás, y de nosotros mismos, que estaba en nuestra mano evitar dicho estado, que podríamos haber hecho más por haber llegado a una concepción más ajustada del asunto que nos ocupaba, más bien que considerarnos víctimas de procesos que dirigen nuestras creencias en sentido contrario a la verdad. Esto no es en sí mismo una razón para rechazar el deflacionismo: es posible que estemos equivocados al considerar que tenemos control sobre nuestro estado en el caso del autoengaño. Pero si disponemos de una explicación que dé cuenta de dicho estado de una manera que concuerde con nuestra idea de control sobre el mismo, esta explicación contaría a primera vista con un dato a su favor, más aún cuando los mismos deflacionistas consideran que controlamos el estado de autoengaño.

Un rechazo del deflacionismo supone recobrar al menos alguno de los elementos que tradicionalmente esta posición considera innecesarios a la hora de dar cuenta del fenómeno del autoengaño. A primera vista es la intencionalidad la que parecería darnos el control que requerimos. Pero el problema al que nos enfrentamos consiste en que si la intención es conocida por el agente, entonces la tarea de engañarse se torna imposible y, como vimos, si es desconocida nos encontramos de nuevo ante una ausencia de control sobre nuestro estado. No es la intención el elemento a recobrar. En mi opinión lo que nos haría responsables de nuestro estado es la posibilidad de atribuir un reconocimiento previo de que la evidencia que el agente tenía y reconocía como tal apuntaba en la dirección contraria a la creencia que finalmente se mantiene. Bermúdez pregunta crucialmente [Bermúdez (2001), p. 316], de dónde proviene la conexión entre nuestro desear que *p* y el que ejercitemos mecanismos distorsionadores para creer que *p*. No todos nuestros deseos desencadenan dichos mecanismos. Ni siquiera la intensidad del deseo puede explicarlo. Sólo un cierto contacto con la verdad del asunto puede desencadenar que la mente

ponga en funcionamiento mecanismos para distorsionarla. Éste no es el problema del carácter selectivo del autoengaño. Este problema aún se suscitaría en el caso de que se considerara que el sujeto estaba en cierto contacto con la verdad como desencadenante del proceso, pues distintos sujetos ante la misma evidencia y con igual deseo podrían reaccionar diferentemente. Se trata del problema de por qué el autoengaño es un caso de engaño y no de simple error. Los ejemplos que Mele proporciona como ejemplos donde los agentes no están en contacto con la verdad parecen funcionar precisamente porque conocemos la verdad y suponemos que ellos también la conocen. ¿Qué diferencia a una madre que falsamente cree que su hijo no consume drogas de otra que se engaña sobre dicho asunto? Sin duda que la segunda dispone de evidencias —las compañías de su hijo, entrevistas con los profesores, los testimonios de vecinos comprometidos, etc.— de las que la primera carece. ¿Qué diferencia al marido que se engaña sobre la infidelidad de su mujer de otro que no lo hace y al que su mujer simplemente engaña? El contacto con la evidencia disponible y el reconocimiento de que la evidencia apuntaba hacia donde apuntaba. Mele argumenta de modo poco convincente que nuestro deseo de que algo sea el caso nos puede llevar a manipular la evidencia para llegar a la creencia contraria a lo que tememos. Así, mi deseo de no ser rechazado por la persona a la que amo me puede llevar recopilar evidencia a favor de la tesis de que no me rechaza. Pero, ¿es esto realmente un caso de autoengaño? Si carece de un modo absoluto de evidencia a favor y en contra acerca de si su amor es correspondido, ¿no le acusaríamos más bien de pensamiento desiderativo? Parece que sólo podríamos acusarle de creer aquello que se desea que sea cierto, pero no de engañarse, pues no había engaño alguno en el proceso de configuración de su creencia, precisamente porque no había reconocimiento alguno del lugar hacia el que apuntaba la evidencia. El único dato para llamar a dicho caso un caso de autoengaño es que, finalmente, la creencia adquirida es falsa y que, a la luz de la evidencia disponible *posteriormente* al proceso de formación de la creencia, se juzga que ese proceso ha conducido a una falsedad⁷. En casos genuinos de autoengaño es el proceso, independientemente de la creencia final resultante, lo que se juzga como un caso de autoengaño debido al reconocimiento de la verdad que, durante el proceso, es atribuible al agente.

Este reconocimiento de que la evidencia apuntaba hacia un lugar distinto del que finalmente se llega es imprescindible para entender el autoengaño y es el factor clave que puede servir para atribuir control y responsabilidad sobre tal estado. Dejando de lado si los mecanismos distorsionadores de dicha evidencia son conscientemente intencionales, inconscientemente intencionales o simplemente no intencionales, ese contacto del sujeto con la evidencia y su reconocimiento en un momento dado del proceso de que la evidencia apuntaba en cierta dirección es suficiente para reprocharle la adquisición de la creencia contraria. De hecho, nosotros mismos nos lo reprocha-

mos cuando la realidad no nos permite retorcer la verdad un poco más. La madre del muchacho que se droga tiene la sensación de que apartó la mirada y no quiso ver lo que sucedía cuando la verdad estaba frente a ella. El marido engañado se siente estúpido porque todo el mundo leía en los datos de los que él también disponía una infidelidad que él no quería ver. El autoengaño vendría precisamente provocado por dicho reconocimiento, y su culpabilidad radicaría en haberse dejado arrastrar por mecanismos de los que se debería haber sospechado por el solo hecho de que hubo un reconocimiento previo de la verdad del asunto en cuestión.

Desde luego que, para que se produzca el autoengaño, es necesario que la evidencia de la que dispone quien se engaña no sea concluyente, al menos desde su propio punto de vista. El grado de responsabilidad variará de acuerdo con qué cantidad y cuán en contacto con la evidencia podamos acusar a alguien de haber estado en un caso de autoengaño. Esta variación en la culpa podría verse acompañada, incluso, de una cierta presunción de intencionalidad en el autoengaño. Así, no necesitamos atribuir a alguien intencionalidad en su estado cuando es poca la evidencia con la que contaba y acepta, al discutir con los demás, la falsedad de la creencia que mantiene. Resulta sin embargo difícil de explicar sin atribuir intencionalidad a su estado el que le hayamos mostrado a alguien la evidencia disponible y hacia donde apunta y que dicha persona continúe empeñada en mantener como cierta la creencia falsa que sostiene. En muchos casos en los que acusamos a alguien de engañarse éste llega a tal estado después de entablar un debate con nosotros en el que va tejiendo razones para defender lo indefendible, para encontrar excusas en un caso particular de lo que, de no ser *su* caso, seguro que consideraría de un modo distinto⁸. No obstante, esta atribución de intencionalidad no es lo más importante para imputarle responsabilidad a alguien; es más bien un indicador de ella. El asunto importante para que haya control del estado de autoengaño es que, de un modo u otro, la persona reconociera en el proceso del autoengaño cuál era la verdad del asunto. Puede que la mente cuente con mecanismos que se activen de un modo no intencionado para manipular la información con la que se cuenta, pero es precisamente el que hayamos estado expuestos a esa información y el que hayamos reconocido que la verdad iba por otros derroteros lo que nos convierte en responsables de nuestro estado.

IV. CONCLUSIÓN

Tal como se ha mantenido a lo largo de este trabajo, el autoengaño es un fenómeno ampliamente extendido. Nos explica cómo alguien puede haber llegado a una creencia manifiestamente falsa y nos sirve en nuestras discusiones cotidianas para acusar a los demás de no querer ver o admitir una verdad que, a todas luces, parece obvia. El problema fundamental reside en

cómo explicar el fenómeno de tal modo que no lo convierta ni en un imposible ni, como en el caso de Mele, en algo tan sencillo e involuntario que haga de nosotros marionetas irresponsables ante el mismo. Mele intenta mostrar que la sola motivación del sujeto puede dar lugar al proceso, no siendo preciso reconocimiento alguno de la verdad por parte del agente para llegar al estado de autoengaño. Mele olvida que el autoengaño no es sólo un estado, sino un proceso en el que siempre existe una tensión entre el reconocimiento de que la evidencia apunta hacia cierto lugar y nuestro deseo de que no sea así. El autoengaño tiene un punto final en el que la persona en cuestión termina creyendo sinceramente una falsedad que no tiene por qué convivir con el reconocimiento previo de la verdad contraria. Podríamos decir que, como en el caso de la muerte, cuando la creencia falsa esté no estará la verdadera, y cuando ésta esté, no lo hará la falsa. Pero esto no quiere decir que no haya habido reconocimiento previo: la reconstrucción final del caso puede apelar a él y uno puede verse acusado de cobardía intelectual, de no querer afrontar una verdad que se nos presentaba de un modo más o menos claro ante nuestros ojos.

Que la mente pueda tener mecanismos no intencionales, innatos o adquiridos socialmente, que nos permitan, como afirmó Nietzsche [Nietzsche (1873), p. 18], “mantenernos un segundo más en la existencia”, es algo que sin lugar a dudas favorece y hace posible el autoengaño. Un control de estos procesos seguramente nos convertiría en agentes epistémicos más responsables, pero no es este punto la clave en la responsabilidad atribuida al que se engaña: la clave sigue estando en un reconocimiento de la verdad que es el que ha generado los procesos distorsionadores de la misma. Sin este reconocimiento previo el autoengaño se convertiría en un proceso que amenazaría cada una de nuestras creencias y, con ellas, la idea de responsabilidad epistémica que pensamos tener.

IES Manuel Tàrraga Escribano
C/ Sancho Panza, s/n
San Pedro del Pinatar, E-30740 Murcia
E-mail: jesuscoll@hotmail.com

NOTAS

¹ No obstante, el concepto de autoengaño no es sólo una explicación para alguna de nuestras creencias, sino que también es utilizado como una acusación antes incluso de que se confirme que es falso lo que el sujeto del autoengaño cree. Esto es importante pues, como se verá, en el análisis de Alfred Mele, el autoengaño parece más bien un estado que un proceso y esto le hace perder de vista que el autoengaño no requiera de un reconocimiento previo de la verdad.

² Véase Davidson (1985), pp. 111-112, para una definición en estos términos.

³ Levy [Levy (2004), p. 297] mantiene que este hecho no tiene por qué beneficiar a los intencionalistas pues, en los casos de autoengaño, la persona que sufre el engaño no tiene por qué estar en mejor situación que otros para interpretar su estado y no hemos de presumir la autoridad de la primera persona. Levy olvida, sin embargo, que el diagnóstico de la primera persona suele coincidir siempre o al menos muy a menudo con el de los demás, quedando así reforzada la idea de que conocíamos de algún modo la verdad.

⁴ Los casos *retorcidos* de autoengaño son aquellos en los que uno termina creyendo algo que no desea que sea el caso. El caso típico es el celoso compulsivo.

⁵ Mele no niega que pueda haber intenciones inconscientes, pero cree que la imagen de la mente que se sigue de dicha suposición —un conjunto de voes mutuamente ocultos unos a otros— le parece lo suficientemente negativa como para abandonar tal hipótesis a favor de una posición no intencionalista.

⁶ Levy pone como ejemplo [Levy (2004), p. 306] que un historiador del Holocausto no tiene la obligación de leer todo el material revisionista que aparece, ni un hombre o una mujer la obligación de buscar continuamente evidencia de que sus parejas les son fieles.

⁷ Se podría pensar que, aun cuando se trate de un caso de pensamiento desiderativo, se trata finalmente, como piensa Mele, de un caso en el que *sólo* la motivación es la que pone en marcha los mecanismos distorsionados de la creencia. En mi opinión esto no es así. Es el desconocimiento de que *p*, junto con el deseo de que *p*, lo que desencadena el que finalmente se crea que *p*. Tanto en el pensamiento desiderativo como en el caso del autoengaño hay un vicio epistémico porque, en el primer caso, sabiendo que no se sabe que *p* se termina creyendo que *p*, mientras que, en el segundo caso, sabiendo o teniendo evidencias de que no *p*, se termina creyendo que *p*. Ambos casos pueden tener fronteras difusas, pues alguien que cree lo que desea puede convertir su caso en un caso de autoengaño si va dejando pasar, o leyendo tendenciosamente la evidencia contraria, precisamente porque es contraria.

⁸ Marcia Baron [Baron (1988), p. 442] afirma que es el hecho de que podamos acudir a diferentes tipos de interpretaciones o historias de una serie de eventos de nuestras vidas lo que sirve de base para que el fenómeno del autoengaño esté ampliamente extendido, e incluso sea algo sin lo cual no podríamos vivir. Esto no significa, en cualquier caso, que el autoengaño sea algo de lo que no seamos responsables. Bien al contrario, nuestra propensión al mismo debe hacer que seamos más cautos en las historias que seleccionamos y mantenemos acerca de nosotros mismos.

REFERENCIAS

- BARNES, A. (1997), *Seeing Through Self-Deception*, Cambridge, Cambridge University Press.
- BARON, M. (1988), "What is Wrong with Self-Deception?", en McLaughlin, P. y Rorty, A. (eds.), *Perspectives on Self-Deception*, Berkeley, California University Press, pp. 431-449.
- BERMÚDEZ, J.L., (2001), "Self-Deception, Intentions and Contradictory Beliefs", *Analysis* 60 (4), pp. 309-319.

- DAVIDSON, D. (1982), "Paradoxes of Irrationality" en Wollheim, R. and Hopkins, J. (eds.), *Philosophical Essays on Freud*, Cambridge, Cambridge University Press, pp. 289-305.
- (1985), "Deception and Division" LePore, E. and McLaughlin, B. (eds.), *Actions and Events*, New York, Basil Blackwell, pp. 138-148.
- JOHNSTON, M. (1988), "Self-Deception and the Nature of Mind", en McLaughlin, P. y Rorty, A. (eds.), *Perspectives on Self-Deception*, pp. 63-91.
- LEVY, N. (2004), "Self-Deception and Moral Responsibility", *Ratio* XVII, pp. 294-311.
- MELE, A. (1987), *Irrationality*, New York, Oxford University Press.
- (1997), "Real Self-Deception", *Behavioral and Brain Sciences* 20, pp. 91-102.
- (2001), *Self-Deception Unmasked*, Princeton, Princeton University Press.
- NIETZSCHE, F. (1873), *Sobre verdad y mentira en sentido extramoral*, edición en castellano de L.M. Valdés, Madrid, Tecnos, 1990.
- RORTY, A., (1988), "The Deceptive Self: Liars, Layers, and Lairs", en McLaughlin, P. y Rorty, A. (eds.), *Perspectives on Self-Deception*, pp. 11-28.