

Autoengaño

1. Caracterización

Una forma de comenzar el examen del autoengaño consiste en describir algunas observaciones preteóricas que resultan intuitivamente plausibles para luego examinarlas en detalle, estrategia que emplearemos aquí. En ocasiones, observamos ciertas conductas (verbales y no verbales) que nos hacen concluir que las personas que las ejecutan poseen una cierta representación falsa acerca de algún aspecto de sí mismas o del mundo externo; esto es, cuando afirman creer algo no están simulando o mintiendo para engañar a otros, sino que parecen actuar de buena fe. Por otra parte, la posesión de dicha representación falsa acerca de algún aspecto de sí mismas o del mundo externo no está apoyada por los elementos de juicio disponibles para la persona, sino que, por el contrario, parece haber pruebas que sostienen la representación opuesta de la que la persona dice poseer; además, parecería que tales elementos de juicio contrarios o bien son poseídos por la persona o bien están fácilmente disponibles para ella. Por último, tenemos razones para pensar que la adopción de tal representación falsa está *motivada*: nos parece que la adopción de la creencia por parte de la persona no es causada por un mero error intelectual generado por factores como la ignorancia, el cansancio o el error de cálculo, sino que se debe a la presencia de estados o procesos no cognitivos que poseen una valencia afectiva (como deseos y temores) que la conducen a aceptar la creencia en cuestión y no su opuesta. Estados como éstos son captados por algunas expresiones de los lenguajes naturales; en el caso de nuestro idioma, por afirmaciones como “se miente a sí mismo” o “está negando la realidad”. Ejemplos típicos de tales procesos son el de la esposa convencida de la fidelidad de su esposo pese a las visibles señales de que tiene un amorío, o el del padre que niega contra todas las pruebas disponibles que su hijo esté sufriendo de un trastorno debido al uso de drogas. Sobre la base de las intuiciones anteriores es posible proponer, entonces, algunos rasgos que parecen caracterizar al autoengaño:

1. adquisición o mantenimiento de una creencia falsa,
2. frente a elementos de juicio contrarios a tal creencia,
3. motivados por procesos mentales no cognitivos (deseos o emociones) que favorecen la adquisición y/o retención de esa creencia.

Así caracterizado, el autoengaño constituye un fenómeno claramente irracional, dado que la persona cree algo incompatible con lo que debería creer dadas las pruebas disponibles y ciertas

normas aceptables de formación de creencias. Pero, debido a que esta irracionalidad pertenece a una clase especial, el autoengaño forma parte de una familia de fenómenos a la que a menudo se hace referencia en el ámbito filosófico con la expresión “irracionalidad motivada”. El término “motivada” es clave para caracterizarla: no se trata de meros sesgos implícitos en el procesamiento de la información, de errores causados por factores como el cansancio o estados mentales alterados, o de conductas impulsivas o descuidadas, sino que estados afectivos como deseos y temores juegan un rol esencial en su producción. Esta familia de fenómenos incluye miembros bien conocidos, como la debilidad de la voluntad (Davidson, 1970) y el pensamiento desiderativo (Szabados, 1973; Bach, 1981). El primero de estos fenómenos consiste en el accionar de una persona en contra de su mejor juicio; es decir, dicha persona evalúa que un curso de acción es superior a sus alternativas en todo respecto y, pese a ello, adopta alguna de esas alternativas. Por su parte, quien piensa desiderativamente adquiere la creencia de que p meramente porque desea creer que es el caso que p .

Ahora bien, para la comprensión del autoengaño resulta de mayor interés el examen del pensamiento desiderativo que la debilidad de la voluntad. Esto se debe a que mientras que el resultado de ésta consiste en una intención o una acción intencional, el resultado del autoengaño consiste en una creencia. Si bien ambos constituyen formas de irracionalidad, el autoengaño estaría incluido en la categoría de irracionalidad teórica, mientras que la debilidad de la voluntad constituiría una forma de irracionalidad práctica. Y, a diferencia de la debilidad de la voluntad, el pensamiento desiderativo también es una forma de irracionalidad teórica. Si bien el rol motivador del deseo caracteriza no sólo al pensamiento desiderativo sino también a la gran mayoría de los casos de autoengaño, ambos parecen diferir en al menos dos aspectos. En primer lugar, el autoengaño parece requerir de una relación entre la persona y la evidencia mucho más compleja que la que caracteriza al pensamiento desiderativo; en segundo lugar, tenemos los llamados casos de “autoengaño retorcido” [*twisted self-deception*], en los cuales la creencia parece ser *contraria* a los deseos de la persona, situación que no tiene paralelo, en principio, en el pensamiento desiderativo. Se han hecho varios intentos de establecer una distinción clara entre ambos fenómenos (Szabados, 1973; Bach, 1981; Davidson, 1985), pero algunos autores (Mele, 1987) no consideran que puedan ser nítidamente diferenciados.

2. El autoengaño en distintas disciplinas

Como ha ocurrido con un gran número de los problemas que hoy podemos encontrar en el campo de la ciencia, el estudio del autoengaño fue planteado originalmente en el ámbito de la filosofía, y de él se han ocupado distintas ramas de esta disciplina: la teoría del conocimiento, la filosofía

moral y, posteriormente, la psicología filosófica y la filosofía de la mente. Posiblemente la referencia más antigua a este fenómeno es debida a Platón en el *Crátilo*, a quien siguieron, entre otros, el obispo Joseph Butler, Immanuel Kant, y posteriormente Karl Marx, Sigmund Freud y Jean-Paul Sartre. Cada uno de ellos se ocupó, por medio de distintas categorías conceptuales y debido a intereses intelectuales muy diversos, de fenómenos muy cercanos a lo que actualmente consideramos casos de autoengaño. A mediados del siglo pasado, el problema comenzó a concitar de modo creciente el interés de distintas ciencias, y en las últimas décadas las investigaciones empíricas sobre el autoengaño se han multiplicado en distintas ramas de la psicología (psicología cognitiva, social, de la personalidad y psicopatología), neuropsicología, teoría de la evolución y ciencias sociales. Estas investigaciones han tenido objetivos muy diversos, desde intentar comprobar empíricamente la existencia del autoengaño (Gur & Sackeim, 1979; Quattrone & Tversky, 1985), hasta determinar los nexos entre el autoengaño y estados patológicos, como los delirios (Bayne & Fernández, 2010). Entre las preguntas empíricas sobre el fenómeno, dos han sido particularmente influyentes. La primera, relativa a la medida en que el autoengaño puede constituir una adaptación en términos evolutivos, esto es, un rasgo de un organismo que contribuye con su capacidad para producir descendencia y perpetuar su patrimonio genético (Von Hippel & Trivers, 2011) o si, por el contrario, se trata de un subproducto estructural, esto es, de un rasgo no seleccionado pero asociado con otros rasgos que sí han sido seleccionados por tal contribución (Van Leeuwen, 2008). La segunda pregunta, a la que volveremos en el penúltimo apartado, refiere a si el autoengaño puede contribuir al bienestar psíquico y/o felicidad (Taylor & Brown, 1988) o si, por el contrario, el bienestar y felicidad obtenidos por medio del autoengaño, aun cuando fuesen posibles, serían ilusorios o efímeros (Colvin, Block & Funder, 1995; Van Leeuwen, 2009).

3. Problemas filosóficos sobre el autoengaño

Pese al desarrollo de influyentes perspectivas científicas sobre diversos aspectos del autoengaño, la persistencia del problema en el campo de la filosofía sigue siendo notoria. Dentro de esta disciplina, podemos distinguir tres grupos de autores. Uno de ellos reúne a las perspectivas que denominaremos “escépticas” (Paluch, 1967; Kipp, 1980; Borge, 2003). Estos filósofos esgrimen diversas razones a favor del escepticismo respecto de la existencia del autoengaño; pero el argumento al que apelan con más frecuencia consiste en resaltar el carácter presuntamente insoluble de las principales paradojas que este fenómeno suscita, a las cuales haremos referencia enseguida. Una segunda posición -estrechamente ligada al construccionismo social y al posmodernismo- es la que podría denominarse “disolucionista” respecto del autoengaño (Harré, 1988; Lewis, 1996; Clegg & Moissinac, 2005). Quienes sostienen esta concepción defienden la tesis según la cual el autoengaño es simplemente una cuestión de perspectivas diferentes, un fenómeno cuya existencia se debe no a procesos intrapsíquicos en el presunto autoengañado sino,

más bien, a la manera en que el observador construye la situación. Finalmente, una abrumadora mayoría de especialistas considera que el autoengaño es un fenómeno genuino que requiere explicación. Entre los filósofos que integran este último grupo, se encuentran Demos (1960), Fingarette (1969), Davidson (1985), Mele (1987, 1997, 2001), Johnston (1988), Oksenberg-Rorty (1988), Talbot (1995), Barnes (1997), Lazar (1999), Funkhouser (2005), Van Leeuwen (2007, 2008, 2009), y muchos otros.

Una buena parte de los esfuerzos filosóficos, consecuentemente, ha estado destinada a explicar cómo sería posible el autoengaño. Los modelos denominados “intencionalistas”, que han constituido durante muchos años la posición estándar sobre el problema (Davidson, 1985; Oksenberg-Rorty, 1988; Pears, 1991), defienden la tesis de que el autoengaño es un fenómeno intencional; esto es, lo que ocurre cuando alguien se engaña a sí mismo es estructuralmente similar a lo que ocurre cuando una persona engaña intencionalmente a otra. Este enfoque, en consecuencia, concibe a la persona autoengañada como alguien que actúa con la intención de que su acción cause la formación de la creencia deseada; tal acción puede consistir en un direccionamiento intencional de la atención lejos de las pruebas que apoyan a p (la creencia que rechaza), o puede consistir en una búsqueda activa de elementos de juicio contra p . Esto no implicaría una intención clara y consciente por parte de la persona de cambiar una creencia verdadera por una falsa, intento que, en principio, se derrotaría a sí mismo. El componente intencional del autoengaño, entonces, debe ser indirecto u oblicuo. A la dificultad para explicar, desde una perspectiva intencionalista, cómo puede tener lugar un proceso intencional de autoengaño se le suele denominar “paradoja dinámica”. Los modelos no intencionalistas o “deflacionistas”, por el contrario, sostienen que el autoengaño es una clase de fenómeno no intencional (Barnes, 1997; Lazar, 1999; Mele, 2001). El autoengaño se explicaría, entonces, por una combinación de procesos cognitivos y factores motivacionales y afectivos. Mele (2001), en particular, es quien más ha desarrollado esta alternativa. El conocimiento acerca de procesos bien estudiados de sesgo cognitivo (de confirmación, heurística de la disponibilidad, saliencia de la información, etcétera), de algunos modelos de testeo de hipótesis cotidianas, y del carácter motivacional de deseos y emociones, permitiría explicar de modo satisfactorio el autoengaño sin atribuirle un carácter intencional. En particular, los sesgos “fríos” o no motivados, si bien pueden actuar independientemente de la motivación, también pueden ser provocados y sostenidos por ésta en la producción de creencias particulares. Así, por ejemplo, la motivación puede incrementar la vivacidad o saliencia de ciertos datos; los datos que avalan la verdad de una hipótesis que uno quisiera que fuera cierta pueden resultar más vívidos o salientes dado el reconocimiento de la dirección en la que apuntan. De este modo, la combinación de tales procesos podría producir un estado de autoengaño aun cuando ninguna intención hubiera participado en su génesis.

Íntimamente relacionado con las concepciones intencionalistas se encuentra un segundo problema fundamental: ¿implica el autoengaño la coexistencia de creencias contradictorias? Si

este fenómeno puede comprenderse a partir del modelo proporcionado por el engaño interpersonal, entonces parecería que debemos admitir que, como ocurre en aquel, la persona que se autoengaña posee tanto la creencia verdadera, que rechaza, como la creencia falsa, que es la que adquiere y/o mantiene (Demos, 1960; Gur y Sackheim, 1979; Davidson, 1985). Esto, como es dable imaginar, conduciría a una situación cognitiva imposible, es decir, la coexistencia en la conciencia de una contradicción flagrante; a este problema se lo suele denominar “paradoja doxástica”. Los modelos deflacionistas, por el contrario, eligen la alternativa de negar que las creencias contradictorias deban coexistir (Barnes, 1997; Lazar, 1999; Mele, 2001). Esto es, para que exista el autoengaño la persona sólo debe poseer la creencia falsa, en presencia de pruebas favorables a la creencia verdadera. Esto podría ocurrir o bien porque sostuvo al inicio y luego desechó la creencia verdadera en favor de la falsa (con lo cual las creencias contradictorias no llegarían a coexistir), o bien no hace falta que haya desestimado, y ni siquiera poseído, la creencia verdadera, sino que adopta inicialmente la creencia falsa. Sean o no admisibles estas alternativas, la posible existencia de creencias contradictorias ha constituido un tema de debate recurrente en torno del autoengaño.

Directamente relacionada con la tesis de la coexistencia de creencias incompatibles se encuentra un tercer e importante interrogante: ¿requiere el autoengaño de una división de la mente? La respuesta a esta pregunta opone a aquellos que han sostenido que tal división es necesaria (Demos, 1960; Davidson, 1985; Oksenberg-Rorty, 1988; Pears, 1991), de aquellos que sostienen que no lo es (Talbot, 1995; Barnes, 1997; Mele, 2001). Quienes responden afirmativamente suelen recurrir a la tesis de una división de la mente como forma de solucionar el problema generado por la presunta coexistencia de creencias contradictorias. Ahora bien, uno de los principales inconvenientes que presentan estas propuestas radica en el carácter *ad hoc* de las divisiones que postulan. Además, dichas divisiones –precisamente en virtud de su naturaleza artificial– generan una imagen inadecuada de la mente; como observa Johnston (1988), éstas parecen ser más apropiadas para dar cuenta de casos de personalidad múltiple que de autoengaño. Por último, algunas de las propuestas divisionistas suponen un compromiso con tesis controvertidas, tales como ideas psicoanalíticas de dudosa científicidad. Por el contrario, las teorías no divisionistas parecen tener a su favor el ser más parsimoniosas, al no proponer ninguna división adicional de la mente a aquellas requeridas para explicar otros fenómenos mentales distintos del autoengaño. No obstante, estas teorías no están exentas de objeciones. Se ha señalado, en particular, que la tensión psíquica que con frecuencia acompaña al autoengaño puede ser mejor explicada si se supone que efectivamente hay creencias contradictorias (o, al menos, representaciones incompatibles de alguna clase; *cfr.* Audi, 1997).

Además de las paradojas doxástica y dinámica, otro problema que aqueja a los modelos intencionalistas es el relativo a la explicación del autoengaño retorcido. En los casos habituales de autoengaño (casos denominados “autoengaño directo”), la persona desea que *p* y, consecuentemente, adquiere o mantiene la creencia de que *p*; en cambio, en los casos de

autoengaño retorcido, la persona adquiere o mantiene la creencia de que $\neg p$, creencia en principio indeseable. Los autores intencionalistas defienden la tesis según la cual la formación o sostenimiento de la creencia autoengañososa se debe siempre a un deseo de adoptarla o mantenerla; por esta razón, el autoengaño retorcido genera un problema para el intencionalismo (Lazar, 1999). Los intencionalistas pueden apelar a dos estrategias para hacer frente a este inconveniente: pueden insistir en que, aun en los casos de autoengaño retorcido, la creencia autoengañososa responde a un deseo de creer, o pueden sostener que el autoengaño retorcido no es genuinamente una variante de autoengaño, sino un fenómeno de otra índole. Frente a la primera réplica, Lazar objeta que su aplicabilidad es limitada: no puede negarse que haya ciertos casos de autoengaño retorcido en los que efectivamente la persona desea adoptar la creencia en juego; sin embargo, instancias de esa índole son, en todo caso, excepcionales. Lazar enfrenta la segunda réplica de los intencionalistas remarcando las considerables similitudes que existen entre el autoengaño directo y el retorcido: ambos tipos suponen la formación o mantenimiento de una creencia que se encuentra a la vez reñida con la evidencia y relacionada temáticamente con factores motivacionales de la persona. Dado el peso de esta convergencia, pierde plausibilidad la idea de que estos fenómenos no puedan incluirse dentro de una misma categoría.

Como puede apreciarse a partir de lo recapitulado hasta aquí, pese a los ingentes esfuerzos teóricos realizados ni intencionalistas ni deflacionistas han logrado presentar explicaciones del fenómeno exentas de problemas más o menos serios (Fernández Acevedo, 2018). Los intencionalistas no logran, en opinión de sus críticos, responder a las paradojas dinámica y doxástica sin generar problemas quizás más graves que los que intentan resolver; estos autores tampoco abordan satisfactoriamente el autoengaño retorcido. Los deflacionistas, por su parte, enfrentan la dificultad de explicar por qué, en determinadas situaciones en las que se cumplen las condiciones requeridas por ellos, el autoengaño no se produce; éste es el denominado “problema de la selectividad” (Bermúdez, 2000). Además, el deflacionismo no logra dar cuenta de un rasgo que para muchos autores (Audi, 1997; Bach, 1997; Losonsky, 1997; Funkhouser, 2005, 2009) es característico del fenómeno: la existencia de una tensión inherente a tal estado, que se manifiesta tanto en la esfera comportamental como en la cognitiva/emocional. La tensión comportamental se expresa en el hecho de que quien se encuentra autoengañado exhibe tanto comportamientos más bien compatibles con la creencia de que p como comportamientos más bien compatibles con la creencia de que $\neg p$; la tensión cognitiva/emocional se advierte en el hecho de que quien se encuentra autoengañado experimenta dudas e inquietudes recurrentes respecto del tópico en juego.

En parte motivadas por las dificultades anteriores, recientemente se han desarrollado algunas propuestas a las cuales, siguiendo a Deweese-Boyd (2016), denominaremos “revisionistas”. El denominador común de tales propuestas consiste en cuestionar uno de los supuestos compartidos por intencionalistas y deflacionistas, a saber, que el producto del autoengaño es una creencia. Una de las principales motivaciones de los revisionistas radica en evitar el surgimiento de la

paradoja doxástica sin renunciar a la intuición de que el autoengaño es un fenómeno signado por una tensión cognitiva/emocional y comportamental. A esta motivación se añade, en algunos casos, el interés intrínseco que reviste considerar el rol que actitudes proposicionales diferentes de las creencias pueden jugar en fenómenos como el autoengaño. Más allá de estas convergencias, prolifera la diversidad entre los autores revisionistas. Así, Gendler (2007) asevera que quien se encuentra autoengañado imagina que p y cree que $\neg p$; Egan (2010) afirma que el producto del autoengaño consiste en una actitud proposicional “intermedia”, que exhibe características distintivas tanto de los deseos como de las creencias; Archer (2013) defiende la tesis según la cual quien se autoengaña no adopta la creencia de que p ni la creencia de que $\neg p$; adopta, en cambio, otras actitudes proposicionales (tales como la sospecha, la duda y la esperanza) respecto de esos contenidos proposicionales. Finalmente, Funkhouser (2009) y Porcher (2012) no rechazan la tesis de que el producto del autoengaño consista en una creencia, pero se apartan tanto de los intencionalistas como de los deflacionistas en cuanto a la manera en que conceptualizan dicha actitud proposicional; basándose en la concepción disposicionalista de Schwitzgebel (2001), argumentan que el producto del autoengaño es una “creencia intermedia” [*in-between belief*]. Las propuestas revisionistas han contribuido al desarrollo de los debates acerca del autoengaño; no obstante, dado su carácter reciente y minoritario, aún es necesario analizar de modo sistemático estas propuestas, a fin de evaluar qué ganancias explicativas aportan (en caso de que lo hagan) respecto de los modelos intencionalistas y deflacionistas.

4. Dimensiones prácticas del autoengaño

Hasta aquí nos hemos ocupado de diversas cuestiones teóricas relativas al autoengaño. En este último apartado, en cambio, examinaremos brevemente algunos aspectos prácticos de este fenómeno (es decir, algunas de las implicaciones y consecuencias que el autoengaño tiene para diversas facetas de nuestras vidas).

Un interrogante de gran relevancia en este ámbito es el siguiente: ¿es quien se autoengaña responsable moralmente por su estado? Como es fácil de imaginar, esta pregunta ha sido una de las de mayor interés dentro del ámbito filosófico, y sin duda es aquella a la que se le ha dado la respuesta que goza del mayor consenso. Este consenso, prácticamente unanimidad, ha consistido en la afirmación de que quien se autoengaña es responsable y censurable por su estado; es digno de reproche desde el punto de vista moral (Baron, 1988; Mele, 2001; Deweese-Boyd, 2008). Como es de imaginar, existe una conexión muy directa entre las concepciones intencionalistas del autoengaño y la posibilidad de atribuir responsabilidad moral: si alguien se autoengaña intencionalmente (aun cuando sea de manera oblicua e indirecta), entonces tiene control sobre

sus acciones y, en consecuencia, es responsable por ellas. Podría pensarse que esta atribución de responsabilidad moral debería ser cuestionada a partir de los enfoques deflacionistas del autoengaño: si este fenómeno no es producto de alguna clase de intención, sino de una compleja combinación de sesgos cognitivos no conscientes y procesos emocionales y motivacionales, parece mucho más difícil adjudicar responsabilidad moral. No obstante, y salvo contadas excepciones (Levy, 2004), casi todos los deflacionistas han defendido la tesis de que quienes se autoengañan son moralmente responsables por su estado. Mele (2001), en particular, señala que quienes se autoengañan son típicamente responsables debido a que los mecanismos de distorsión que usualmente intervienen (como el sesgo de confirmación) están en alguna medida bajo nuestro control; si conocemos su existencia, podemos instruirnos a nosotros mismos de modo de minimizar sus efectos. Por supuesto, no podemos librarnos por completo de su actuación, pero cuando hay cuestiones importantes en juego podemos recordarnos a nosotros mismos de su potencial de sesgo y actuar en consecuencia. Cabe agregar, por último, que los deflacionistas podrían recurrir al argumento de que lo que es un *modus ponens* para un filósofo es un *modus tollens* para otro: si la corrección de su modelo implica el abandono del supuesto de intencionalidad y, consecuentemente, de responsabilidad moral, entonces quien se autoengaña no es moralmente responsable por su estado.

Otro aspecto práctico del autoengaño que ha suscitado gran interés es el relativo al impacto de éste sobre la felicidad y la salud mental. En lo que concierne a esta cuestión, cabe señalar que si bien el rol potencialmente defensivo del autoengaño contra verdades dolorosas fue señalado reiteradamente dentro del ámbito filosófico (Williams, 1973; Davidson, 1985), no fue sino hasta la publicación de diversos trabajos de la psicóloga Shelley Taylor y sus colaboradores (Taylor & Brown, 1988, 1994; Taylor, Kemeny, Reed, Bower & Gruenewald, 2000) que sus posibles beneficios comenzaron a ser discutidos de manera sistemática, tanto dentro de la psicología como de la filosofía. En su muy debatido artículo de 1988, Taylor y Brown sostuvieron que ciertas formas “suaves” de autoengaño harían posible, para quienes las experimentan, una mayor felicidad y salud mental en comparación con aquellas personas cuyas visiones de la realidad externa y de sí mismos son más precisas. Estas tesis, además de ser cuestionadas desde la propia psicología (Colvin & Block, 1994; Colvin, Block & Funder, 1995) fueron objetadas por diversos filósofos (Bhadwar, 2008; Van Leeuwen, 2009; Martin, 2012), quienes señalaron los riesgos de suponer que el autoengaño podría constituir la base para el logro de una felicidad genuina. Más aun, este último autor sostuvo que el autoengaño, en lugar de constituir un fundamento sólido para una forma de felicidad digna de ese nombre, sería en realidad un obstáculo para ella.

La última cuestión que examinaremos es la relativa a la medida en que el autoengaño se encuentra presente en sistemas de creencias colectivos de distintas clases (Broad & Wade, 1982; Goleman, 1989; Deweese-Boyd, 2008). Dichos sistemas de creencias revisten clara significatividad práctica, puesto que refieren a cuestiones de relevancia social, política, económica o cultural. Así, por mencionar un ejemplo, Rääkkä (2007) y Lynch (2010) han

examinado el autoengaño colectivo vinculado con creencias de índole religiosa, de mucha importancia práctica para la vida de diversas comunidades. La posibilidad del autoengaño colectivo está ligada a diversos debates referidos a la forma de concebir la atribución de actitudes a grupos o comunidades (Deweese-Boyd, 2008). Con relativa independencia de estos debates, hay razones para pensar que la expresión “autoengaño colectivo” no refiere a una colección de autoengaños individuales reforzados por procesos de interacción social, sino que se trata de un proceso en el cual la dimensión social tiene una relevancia explicativa como mínimo equivalente a la dimensión psicológica. Los procesos de distorsión colectiva de la realidad incluyen interacciones sociales en diversas dimensiones, que involucran procesos de influencia en escala tanto micro y macrosocial como en dirección ascendente y descendente (Fernández Acevedo, 2018).

5. Conclusión

Como puede apreciarse a partir esta sucinta presentación, el autoengaño es un fenómeno que genera numerosas y complejas controversias. A pesar de ello, una idea parece gozar de amplio consenso: el autoengaño está profundamente enraizado en nuestra naturaleza. Cualesquiera sean sus orígenes, sin duda nuestros sistemas perceptivos, cognitivos y afectivos se encuentran estructurados e interrelacionados de un modo tal que hacen posible, o incluso facilitan, el autoengaño, así como otras formas de irracionalidad motivada. Quizás podamos incrementar, tanto de modo individual como colectivo, nuestra capacidad para ser conscientes de su actuación y minimizarla, pero parece improbable (y, tal vez, indeseable) que logremos eliminarlo por completo.

Gustavo Fernández Acevedo
(Universidad Nacional de Mar del Plata)

Agostina Vorano
(Consejo Nacional de Investigaciones Científicas y Tecnológicas, Universidad Nacional de Mar del Plata)

Referencias

- Audi, R. (1997). Self-deception vs. self-caused deception: a comment on Professor Mele. *Behavioral and Brain Sciences*, 20(1), 104.
- Archer, S. (2013). Non-doxasticism about self-deception. *Dialectica*, 67(3), 265-282.
- Bach, K. (1981). An analysis of self-deception. *Philosophy and Phenomenological Research*, 41(3), 351-370.
- Bach, K. (1997). Thinking and believing in self-deception. *Behavioral and Brain Sciences*, 20(1), 105.
- Badhwar, N. (2008). Is realism bad for you? A realistic response. *The Journal of Philosophy*, 105(2), 85-107.
- Barnes, A. (1997). *Seeing through self-deception*. New York, USA: Cambridge University Press.
- Baron, M. (1988). What is wrong with self-deception? In B. P. McLaughlin & A. Oksenberg-Rorty (Eds.), *Perspectives on self-deception* (pp. 431-449). Berkeley, USA: University of California Press.
- Bayne, T. & Fernández, J. (Eds.) (2010) *Delusion and self-deception. Affective and motivational influences on belief formation*. New York, USA: Psychology Press.
- Bermúdez, J. L. (2000). Self-deception, intentions and contradictory beliefs. *Analysis*, 60(4), 309-319.
- Borge, S. (2003). The myth of self-deception. *The Southern Journal of Philosophy*, 41(1), 1-28.
- Broad, W. & Wade, N. (1982). Self-deception and gullibility. In W. Broad & N. Wade (Eds.), *Betrayers of the Truth*. New York, USA: Simon and Schuster.
- Clegg, J. W. & Moissinac, L. (2005). A relational theory of self-deception. *New Ideas in Psychology*, 23(2), 96-110.
- Colvin, R. & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin*, 116(1), 3-20.
- Colvin, R., Block, J. & Funder, D. (1995). Overly positive self-evaluations and personality: negative implications for mental health. *Journal of Personality and Social Psychology*, 68(6), 1152-1162.
- Davidson, D. (1970). ¿Cómo es posible la debilidad de la voluntad? En *Ensayos sobre acciones y sucesos*. México DF, México: UNAM - Crítica.
- Davidson, D. (1985). Engaño y división. En *Mente, mundo y acción*. Barcelona, España: Paidós.
- Demos, R. (1960). Lying to oneself. *Journal of Philosophy*, 57, 588-595.
- Deweese-Boyd, I. (2008). Collective self-deception, collective injustice: consumption, sustainability and responsibility. Extraído el 27/10/10 de http://www.colorado.edu/philosophy/center/rome/papers/DeWeese-boyd_CollectiveSelfDeception_CollectiveInjustice.pdf
- Deweese-Boyd, I. (2016). Self-deception. En E. N. Zalta (Ed.), *Stanford Encyclopaedia of*

- Philosophy*. Extraído el 10/06/17 de. <http://plato.stanford.edu/entries/self-deception/>
- Egan, A. (2010). Imagination, delusion, and self-deception. In T. Bayne & J. Fernández (Eds.), *Delusion and self-deception. Affective and motivational influences on belief formation*. New York, USA: Psychology Press.
 - Fernández Acevedo, G. (2018). *El autoengaño. Anatomía de una pasión humana*. Buenos Aires, Argentina: Sociedad Argentina de Análisis Filosófico.
 - Fingarette, H. (1969). *Self-deception*. London, UK: Routledge & Kegan Paul.
 - Funkhouser, E. (2005). Do the self-deceived get what they want? *Pacific Philosophical Quarterly*, 86(3), 295-312.
 - Funkhouser, E. (2009) Self-Deception and the limits of Folk Psychology. *Social Theory and Practice*, 35(1), 2-13.
 - Gendler, T. S. (2007). Self-deception as pretense. *Philosophical Perspectives*, 21, 231-258.
 - Goleman, D. (1989). What is negative about positive illusions? When benefits for the individual harm the collective. *Journal of Social and Clinical Psychology*, 8, 190-197.
 - Gur, R. C. & Sackeim, H. A. (1979). Self-deception: a concept in search of a phenomenon. *Journal of Personality and Social Psychology*, 37(2), 147-169.
 - Harré, R. (1988). The social context of self-deception. In B. P. McLaughlin & A. Oksenberg-Rorty (Eds.), *Perspectives on self-deception* (pp. 364-379). Berkeley, USA: University of California Press.
 - Johnston, M. (1988). Self-deception and the nature of mind. In B. P. McLaughlin & A. Oksenberg-Rorty (Eds.), *Perspectives on self-deception* (pp. 63-91). Berkeley, USA: University of California Press.
 - Kipp, D. (1980). On self-deception. *Philosophical Quarterly*, 30(121), 305-317.
 - Lazar, A. (1999). Deceiving oneself or self-deceived? On the formation of beliefs 'under the influence'. *Mind*, 108(430), 265-290.
 - Levy, N. (2004). Self-deception and moral responsibility. *Ratio*, 17(3), 294-311.
 - Lewis, B. (1996). Self-deception: a postmodern reflection. *Journal of Theoretical and Philosophical Psychology*, 16(1), 49-66.
 - Losonsky, M. (1997) Self-deceivers' intentions and possessions. *Behavioral and Brain Sciences*, 20(1), 121-122.
 - Lynch, K. (2010). Self-deception, religious belief, and the false belief condition. *The Heythrop Journal*, 51(6), 1073-1074.
 - Martin, M. (2012). *Happiness and the good life*. New York, USA: Oxford University Press.
 - Mele, A. R. (1987). *Irrationality. An essay on akrasia, self-deception, and self-control*. Oxford, UK: Oxford University Press.
 - Mele, A. R. (1997) Real self-deception. *Behavioral and Brain Sciences*, 20, 91-136.
 - Mele, A. R. (2001). *Self-deception unmasked*. Princeton, USA: Princeton University Press.
 - Oksenberg-Rorty, A. (1988). The deceptive self: liars, layers and lairs. In B. P. McLaughlin & A. Oksenberg-Rorty (Eds.), *Perspectives on self-deception* (pp. 11-28). Berkeley, USA: University of California Press.
 - Paluch, S. (1967). Self-deception. *Inquiry*, 10(1-4), 268-278.
 - Pears, D. (1991). Self-deceptive belief-formation. *Synthese*, 89, 393-405.
 - Porcher, J. (2012). Against the deflationary account of self-deception. *Humana.Mente*

Journal of Philosophical Studies, 20, 67-84.

- Quattrone, G. & Tversky, A. (1985). Self-deception and the voter's illusion. In J. Elster (Ed.), *The Multiple Self*. Cambridge, UK: Cambridge University Press.
- Rääkkä, J. (2007). Self-deception and religious beliefs. *The Heythrop Journal*, 48(4), 513-526.
- Schwitzgebel, E. (2001). In-between believing. *The Philosophical Quarterly*, 51(202), 76-82.
- Szabados, B. (1973). Wishful thinking and self-deception. *Analysis*, 33(6), 201-205.
- Talbott, W. (1995). Intentional self-deception in a single coherent self. *Philosophy and Phenomenological Research*, 55(1), 27-74.
- Taylor, S. & Brown, J. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193-210.
- Taylor, S. & Brown, J. (1994). Positive illusions and well-being revisited. Separating fact from fiction. *Psychological Bulletin*, 116(1), 21-27.
- Taylor, S., Kemeny, M., Reed, G., Bower, J. & Gruenewald, T. (2000). Psychological resources, positive illusions, and health. *American Psychologist*, 55(1), 99-109.
- Van Leeuwen, D. S. N. (2007). The product of self-deception. *Erkenntnis*, 67(3), 419-437.
- Van Leeuwen, D. S. N. (2008). Finite rational self-deceivers. *Philosophical Studies*, 139(2), 191-208.
- Van Leeuwen, D. S. N. (2009). Self-deception won't make you happy. *Social Theory and Practice*, 35(1), 107-132.
- von Hippel, W. & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, 34(1), 1-56.
- Williams, B. (1973). Deciding to believe. In *Problems of the self. Philosophical papers 1956-1972* (pp. 136-151). New York, USA: Cambridge University Press.

Lecturas recomendadas en castellano

- Davidson, D. (1985). Engaño y división. En *Mente, mundo y acción*. Barcelona, España: Paidós.
- Gomila, A. (2007). El retorno de la represión. *Teorema*, 26(3), 97-111.
- Saab, S. (2011). Modos de autoengaño y de razonamiento: teorías de proceso dual. *Análisis filosófico*, 31(2), 193-218.
- Trivers, R. (2011). *La insensatez de los necios*. Buenos Aires, Argentina: Katz Editores.

Recursos on-line

- Deweese-Boyd, I. (2016). Self-deception. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/self-deception/>
- Kirsch, J. (s/f). Ethics and self-deception. In J. Fieser & B. Dowden (Eds.), *The Internet Encyclopedia of Philosophy*. <https://www.iep.utm.edu/eth-self/>

Entradas relacionadas

- **Akrasia**
- **Psicología (Filosofía de la)**

Cómo citar esta entrada

Fernández Acevedo, G. & Vorano, A. (2021). Autoengaño. *Enciclopedia de la Sociedad Española de Filosofía Analítica*. (URL: <http://www.sefaweb.es/autoengano/>)