

IV Congreso Internacional de Investigación y Práctica Profesional en Psicología  
XIX Jornadas de Investigación VIII Encuentro de Investigadores en Psicología  
del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos  
Aires, 2012.

# **El autoengaño: la transición de un problema desde la filosofía a la psicología.**

Fernández Acevedo, Gustavo.

Cita:

Fernández Acevedo, Gustavo (2012). *El autoengaño: la transición de un problema desde la filosofía a la psicología*. IV Congreso Internacional de Investigación y Práctica Profesional en Psicología XIX Jornadas de Investigación VIII Encuentro de Investigadores en Psicología del MERCOSUR. Facultad de Psicología - Universidad de Buenos Aires, Buenos Aires.

Dirección estable: <https://www.aacademica.org/000-072/110>

*Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.*

# EL AUTOENGAÑO: LA TRANSICIÓN DE UN PROBLEMA DESDE LA FILOSOFÍA A LA PSICOLOGÍA

Fernández Acevedo, Gustavo

Universidad Nacional de Mar del Plata. Argentina

---

## Resumen

El fenómeno del autoengaño ha formado parte del repertorio de problemas filosóficos durante varios siglos. No obstante, en las últimas tres o cuatro décadas ha comenzado a ser objeto de análisis sistemáticos por parte de diversas disciplinas científicas, entre ellas la psicología. Diversas ramas de esta disciplina, entre las que se cuentan la psicología social, la psicología cognitiva, la psicología de la personalidad y la psicopatología han contribuido a nuestra comprensión de este fenómeno. Sin embargo, la transición del problema desde la filosofía a la psicología no carece de dificultades conceptuales, originadas al menos en parte en el descuido de distinciones filosóficas conocidas. En esta ponencia se examinan dos aspectos problemáticos de esta transición: en primer lugar, el desarrollo de estudios psicológicos sobre el autoengaño, tanto teóricos como empíricos, basados en definiciones conceptuales inadecuadas del fenómeno y, en segundo lugar, la adopción de supuestos filosóficos generales cuestionables (tanto ontológicos como gnoseológicos) para su comprensión.

## Palabras Clave

Autoengaño, Psicología, Filosofía

## Abstract

SELF-DECEPTION: THE TRANSITION OF A PROBLEM FROM PHILOSOPHY TO PSYCHOLOGY

The phenomenon of self-deception has been part of the stock of philosophical problems for many centuries. However, for the last three of four decades this phenomenon has become subject of systematic studies to several scientific disciplines, including psychology. Several brands of this science, including social psychology, cognitive psychology, personality psychology and psychopathology have contributed to our understanding of self-deception. Nevertheless, the transition of this problem from philosophy to psychology does not lack conceptual troubles, whose origins are in the neglect of known philosophical distinctions. In this work two problematic aspects of this transition are examined. First, the development of psychological studies on self-deception, both theoretical and empirical, based on inadequate conceptual definitions of the phenomenon; second, the adoption of disputable philosophical assumptions (both ontological and epistemological) for the understanding of the phenomenon.

## Key Words

Self-deception, Psychology, Philosophy

I

Una de las posiciones más difundidas respecto de la relación entre los problemas científicos y los problemas filosóficos es la que sostiene que los primeros suceden y derivan, conceptual y cronológicamente, de los segundos. Esta tesis ha sido expresada por el filósofo John Austin (1961) de una manera que se ha hecho célebre: "En la historia de las indagaciones humanas la filosofía ocupa el lugar de un sol central originario, seminal y tumultuoso. De tanto en tanto ese sol arroja algún trozo de sí mismo que adquiere el *status* de una ciencia, de un planeta frío y bien regulado, que progresa sin pausa hacia un distante estado final' (p. 180).

El problema del autoengaño ilustra con notable nitidez y actualidad tal proceso de transición por el cual un problema filosófico pasa a ser un problema científico. Diversas disciplinas, entre las que se cuentan la biología evolucionista (Trivers, 2000, 2011), las neurociencias (Ramachandran, 2006; Turnbull y Solms, 2007), y las ciencias sociales (Cohen, 2001; Zerubavel, 2006) se han interesado activamente en los últimos años en la descripción y explicación del fenómeno del autoengaño.

Si bien hay antecedentes bastante remotos de intentos de estudiar empíricamente el fenómeno (Frenkel-Brunswick, 1939), no es sino hasta hace algo más de tres décadas que su examen forma parte de manera más sistemática de la agenda de la investigación psicológica. La psicología cognitiva, la psicología social, la psicología de la personalidad y la psicopatología han hecho aportes para la comprensión de un fenómeno singularmente elusivo y complejo.

El análisis del proceso de transición particular por el cual el autoengaño, sin abandonar por completo su pertenencia al campo de la filosofía, pasa a ser parte de los intereses de investigación de los psicólogos, presenta aristas epistemológicas de interés. Se trata de un proceso complejo, en el que pueden observarse, en nuestra opinión, tanto ganancias como pérdidas. La reflexión sobre este proceso hace posible varias consideraciones potencialmente útiles. En primer lugar, poner de manifiesto lagunas o déficits en la conceptualización psicológica sobre el fenómeno, originadas en ocasiones por una negligencia respecto de teorizaciones filosóficas previas. En segundo lugar, y derivado de lo anterior, posibilita sugerir líneas alternativas y más promisorias de investigación. Permite, por último, renovar la reflexión acerca de las relaciones entre la psicología y la filosofía, relaciones siempre complejas y vigentes pese a la ya remota separación de ambas disciplinas.

Si bien son múltiples las dimensiones epistemológicas susceptibles de exploración respecto de la transformación del problema del autoengaño en un tema de investigación para la psicología, en esta presentación nos limitaremos a presentar dos aspectos problemá-

ticos. Ellos son: los problemas conceptuales en la caracterización del fenómeno, producto en parte del descuido de la teorización filosófica previa, en primer lugar, y la adopción de supuestos filosóficos generales cuestionables para la comprensión del fenómeno, en segundo término.

## II

Uno de los rasgos notables que pueden observarse con frecuencia en los estudios psicológicos sobre el autoengaño es el descuido de precisiones conceptuales y, correlativamente, la adopción de definiciones cuestionables del fenómeno.

Paulhus (2007), desde la psicología social, define al autoengaño como 'el acto de mentirse a uno mismo'. Refiere a esta conducta como casos en los cuales las 'personas aparentemente creen algo que saben que es falso'. Aclara que esta conducta no incluye exageración, falsificación o simple mentira; el autoengaño es algo más profundo y complicado, incluso paradójico. Ejemplos de este fenómeno son los casos de un hombre habitualmente agradable que bebe demasiado y rechaza la creencia de que tiene un problema con el alcohol o la madre de un criminal que rechaza lo que la policía dice de él. Lo que caracteriza a todos estos casos es que la evidencia en favor de la creencia rechazada es mayor que la evidencia contraria. Paulhus considera que la explicación de casos como éstos requiere el reconocimiento de la existencia de partes inconcientes de la mente. Sólo en tales partes inconcientes puede un conflicto emocional influenciar realmente la conducta de un individuo y, pese a eso, ser inaccesible.

Así como a Paulhus la coexistencia de creencias contradictorias y la postulación de una concepción agencial del fenómeno (esto es, las personas somos agentes activos en la producción del autoengaño) les parecen perfectamente aceptables, para otros psicólogos estas mismas tesis son tan evidentemente incorrectas como para rechazar la posibilidad del autoengaño. Bandura (1991) ha negado de manera tajante la existencia de este fenómeno. Basa su posición escéptica en la conocida tesis de la imposibilidad de ser simultáneamente quien engaña y quien es engañado; es lógicamente imposible, señala, engañarse a uno mismo para creer algo mientras simultáneamente se sabe que es falso. Observa, asimismo, que los intentos de resolver esta paradoja mediante la apelación a partes inconcientes de la mente han tenido escaso éxito. Considera que esta estrategia, más que explicar el autoengaño, lo aniquila. Estas concepciones del yo dividido, como las denomina, fracasan al tener que explicar como un yo conciente puede mentir a un yo inconciente; el yo engañador tiene que ser conciente de lo que el yo engañado cree para saber como fraguar el engaño.[i]

Un tercer ejemplo de adopción de definiciones cuestionables se encuentra en los intentos experimentales (por otra parte muy destacados) de probar la existencia del autoengaño. Lugar de privilegio en esta vertiente es ocupado por los estudios de Gur y Sackheim (1979). Estos autores proponen los siguientes criterios como necesarios y suficientes para el autoengaño: '1. el individuo sostiene dos creencias contradictorias (que *p* y que no *p*). 2. Esas dos creencias contradictorias son sostenidas simultáneamente. 3. El individuo no es conciente de que sostiene una de esas creencias. 4. El acto que determina cual creencia es objeto de la conciencia y cual no es un acto motivado' (p. 149).

Los tres estudios reseñados son susceptibles de una misma clase

de objeción: el partir de caracterizaciones como mínimo discutibles del fenómeno del autoengaño. Esto se debe a que todos comparten un mismo supuesto cuestionable: el de que el autoengaño debe implicar esencialmente la coexistencia de creencias contradictorias. La adopción de este supuesto conduce, a su vez, a dos alternativas igualmente innecesarias: o bien, como es el caso de Paulhus y Gur y Sackheim, a proponer partes inconcientes de la mente de una forma *ad hoc*, o bien, como es el caso de Bandura, a negar la existencia del fenómeno.

El supuesto de la coexistencia de creencias contradictorias ha sido cuestionado de manera convincente por muchos estudios filosóficos (Talbot, 1995; Barnes, 1997; Mele, 1987, 2001). Para comenzar, es pertinente observar que las creencias contradictorias podrían no coexistir, sino suceder una a la otra en el tiempo. Sin embargo, una alternativa más radical consiste en negar que la creencia verdadera forme parte inicialmente del repertorio de creencias del agente y que luego sea sustituida por la creencia falsa; esto es, las creencias contradictorias no coexisten ni se suceden temporalmente en la mente del agente. En Mele (2001), por ejemplo, puede encontrarse incluso un modelo explicativo del autoengaño que prescinde de tal supuesto y basado en sólida evidencia empírica provista por varios estudios psicológicos. En esta perspectiva el autoengaño es el producto de una combinación de sesgos cognitivos "fríos" y estados motivacionales y emocionales, que conduce a la adopción de creencias falsas sin que la creencia verdadera o más fundada en la evidencia haya estado alguna vez en posesión del agente.

Encontramos, entonces, que el descuido de análisis conceptuales detallados provenientes de la filosofía conduce a consecuencias teóricas y empíricas inadecuadas.

## III

No han faltado tampoco intentos, desde la psicología, de "disolver" el fenómeno del autoengaño. Estos intentos suelen estar ligados a posiciones construccionistas radicales, que adoptan posiciones subjetivistas y relativistas tanto en el plano ontológico como en el plano gnoseológico. Si bien hay varios escritos en esta vertiente (Gergen, 1985; Lewis, 1986; Harre, 1988), tomaremos como modelo el trabajo de Clegg y Moissinac (2005).

Siguiendo a Gergen, estos autores observan que el término autoengaño posee poder ilocucionario como una herramienta retórica, más que como algo que hace referencia a un fenómeno mental específico. Observan, también basados en Gergen, que existen además dos modos en los cuales el concepto de autoengaño provee valor cultural: posibilita a una persona evitar la responsabilidad por acciones irresponsables y es un modo de ejercer poder sobre otra persona.

Respecto de la explicación del fenómeno, los autores rechazan la noción de procesos psíquicos que sesgan nuestra visión de la realidad. La noción de proceso sesgado supone una clase de realidad externa absoluta contra la cual el procesamiento erróneo puede ser medido -una creencia falsa requiere una verdad, así como los medios para arbitrar entre ellas-. Como todas las ramas del psicologismo, observan, este enfoque corteja el absurdo debido a que las creencias supuestamente verdaderas o garantizadas respecto de las cuales las creencias sesgadas son comparadas, son en sí mismas los productos de alguna clase de razonamiento sesgado.

La denominada “teoría relacional de la conciencia” que sugieren los autores tiene como afirmación principal que los objetos de la experiencia no son reducibles más allá de sus relaciones primarias -relaciones tanto dentro como entre interlocutores-. El mundo objetivo es a la vez único y común -único porque cada ‘objeto’ está irreductiblemente constituido en actos de experiencia particular y común porque otros interlocutores también constituyen cada ‘objeto’. El mundo de los objetos, entonces, no es independiente del discurso o la conciencia’. Este modelo de la conciencia, señalan, implica cosas importantes para la teoría del autoengaño. En primer lugar, implica que todos los ‘objetos’ son idiosincrásicos, esto es, no hay una realidad externa objetiva unificada que confiera autoridad epistémica a un observador externo. Más aun, implica que las discrepancias evidentes en los casos paradigmáticos de autoengaño no son el resultado de una conciencia fragmentada cegada a sí misma a los hechos inevitables de la realidad ‘objetiva’, sino, más bien, el resultado de diferentes observadores que captan diferentes objetos idiosincrásicos, diferentes mundos fenomenológicos y hacen en consecuencia evaluaciones radicalmente diferentes. Los autores presentan como ejemplo de supuesto autoengaño el caso de un padre cuya hija consume drogas y que muere a causa de esta conducta, conducta de la que el padre no parece percatarse. Mientras que las perspectivas “clásicas” hablarían de una distorsión autoengañososa en la percepción del padre, Clegg y Moissinac enfatizan las diferencias en las construcciones idiosincrásicas del padre y de los observadores externos.

Tenemos aquí, entonces, una segunda forma en la que la filosofía permea los estudios psicológicos sobre el autoengaño, esto es, mediante la adopción de una perspectiva filosófica radicalmente constructorista y relativista. El análisis y eventual crítica del constructorismo radical excede ampliamente el marco de este trabajo, por lo que sólo es posible realizar algunas observaciones no sistemáticas. Por otro lado, no es necesario cuestionar aquí la adopción de una filosofía constructorista en su totalidad; basta cuestionar su adopción para el caso particular.

Si bien, como se verá enseguida, algunos psicólogos han señalado que la adopción de creencias falsas sobre uno mismo puede tener eventualmente efectos positivos, esta posibilidad aislada no anula el hecho de que el éxito en nuestro trato con el mundo depende esencialmente de que al menos algunas de nuestras creencias sean verdaderas. Ahora bien, si no hay nada más que construcciones idiosincrásicas del mundo y cualquier construcción es tan válida como cualquier otra, ¿por qué algunas construcciones nos proporcionan un éxito mayor en nuestro trato con el mundo y otras no lo hacen? En el caso particular del ejemplo propuesto por Clegg y Moissinac, se debería explicar por qué la construcción del padre conduce a un resultado totalmente indeseable, cosa que podría no ocurrir, en principio, con la posición de un observador externo que afirme la existencia de una adicción a una droga. La posición filosófica constituida por el constructorismo radical enfrenta, entonces, similares problemas a los que enfrenta en el plano epistemológico: si las teorías científicas son meramente construcciones relativas de idéntico valor epistémico, entonces no es posible explicar por qué algunas de esas teorías nos conducen a un trato exitoso con el mundo, mientras otras no lo hacen.

#### IV

Lo expuesto resulta suficiente, a nuestro modo de ver, como para justificar la afirmación inicial relativa a la complejidad de la transición del autoengaño desde la filosofía a la psicología.

En este análisis hemos dejado de lado, por razones de brevedad, consideraciones acerca de los límites borrosos entre ambas disciplinas en lo que respecta a este problema. Este carácter difuso de los límites disciplinares no sólo se extiende a las estrategias adecuadas para estudiar el problema, sino también a lo que podríamos llamar “división del trabajo” entre especialistas de diferentes disciplinas. Por caso, el citado texto de Mele (2001) ofrece un notable ejemplo de teorización psicológica, basada en evidencia empírica, realizado por un filósofo. Por otro lado, conviene observar que la existencia de una transformación del problema no implica que tal transición haya traído aparejada la pérdida de interés filosófico en el problema. Algunos problemas relativos al autoengaño, como el de la responsabilidad moral por encontrarse en tal estado, parecen seguir siendo típicamente filosóficos.

Las críticas expuestas no implican sustentar la prescripción (imposible de cumplir, por otra parte) de que los estudios psicológicos sobre el autoengaño deben aguardar a que la filosofía solucione las cuestiones conceptuales fundamentales que subyacen al fenómeno para luego estudiarlo de manera empírica. Significa solamente que los psicólogos debemos ser cuidadosos a la hora de examinar y aceptar tesis conceptuales específicas relativas al fenómeno en cuestión.

Los psicólogos, en síntesis, haríamos bien en no perder de vista los aportes filosóficos sobre el problema del autoengaño. Sólo así será posible aproximarnos a la constitución de algo similar al “planeta frío y bien regulado” mencionado en las palabras de Austin.

#### Bibliografía

- Austin, J. (1961). *Philosophical Papers*. G. J. Warnock y J. O. Urmson (eds.), Oxford, Clarendon Press.
- Bandura, A. (1991). *Social cognitive theory of moral thought and action*. En W. M. Kurtines & J. L. Gewirtz (eds.), *Handbook of moral behavior and development*. Vol. 1. Hillsdale, NJ, Erlbaum. 45-103.
- Bandura, A. (2011). *Self-deception: A paradox revisited*. *Behavioral and Brain Sciences*, 34, 16-17.
- Barnes, A. (1997). *Seeing through self-deception*. New York, Cambridge University Press.
- Clegg, Joshua & Luke Moissinac (2005). *A relational theory of self-deception*. *New Ideas in Psychology*, 23. 96-110.
- Cohen, S. (2001). *States of Denial. Knowing about Atrocities and Suffering*. Oxford, Blackwell.
- Frenkel-Brunswick, E. (1939). *Mechanisms of Self-Deception*. *Journal of Social Psychology*, 10.
- Gergen, K. (1985). *The ethnopsychology of self-deception*. In M. W. Martin (Ed.), *Self-deception and self-understanding*. Lawrence, KS: University Press of Kansas.
- Gur, R. & Harold S. (1979). *Self-Deception: A Concept in Search of a Phenomenon*. *Journal of Personality and Social Psychology* 37:147-69.
- Harre, R. (1988). *The Social Context of Self-Deception*. En B. McLaughlin & A. Oksenberg-Rorty (eds.), *Perspectives on Self-Deception*. Berkeley, University of California Press.
- Lewis, B. (1996). *Self-deception: A postmodern reflection*. *Journal of Theoretical and Philosophical Psychology*, 16, 1. 49-66.
- Mele, A. (1987). *Irrationality. An Essay on Akrasia, Self-Deception, and Self-Control*. New York-Oxford, Oxford University Press.
- Mele, A. (2001). *Self-deception Unmasked*. Princeton, Princeton University Press.
- Paulhus, D. (2007). *Self Deception*. En Roy Baumeister & Kathleen Vohs (eds.), *Encyclopedia of Social Psychology*. Los Angeles, Sage.
- Ramachandran, V. S. (1996). *The Evolutionary Biology of Self-Deception, Laughter, Dreaming and Depression: Some Clues for Anosognosia*. *Medical*

Hypothesis 47. 347-362.

Talbott, W. (1995). Intentional Self-Deception in a Single Coherent Self. *Philosophy and Phenomenological Research* 55, 1. 27-74.

Trivers, R. (2000). The Elements of a Scientific Theory of Self-Deception. *Annals of the New York Academy of Sciences* 907. 114-131.

Trivers, R. (2011). *The Folly of Fools: The Logic of Deceit and Self-Deception in Human Life*. New York, Basic Books.

Turnbull, O. & Mark S. (2007). Awareness, Desire, and False Beliefs: Freud in the Light of Modern Neuropsychology. *Cortex* 43. 1083-1090.

Zerubavel, E. (2006). *The Elephant in the Room. Silence and Denial in Everyday Life*. New York, Oxford University Press.