

8

Identidad y autoengaño en la Cuarta Revolución*

Carlos Patarroyo G.**

* Este texto ha sido posible gracias al apoyo del Ministerio de Economía y Competitividad de España, mediante la financiación concedida al proyecto de investigación titulado: “Autoconocimiento, responsabilidad moral y autenticidad” (FFI2016-75323-P).

** Decano, Escuela Ciencias Humanas, Universidad del Rosario. Ph. D. en Filosofía.

*Tenemos, quienes vivimos,
una vida que es vivida
y otra vida que es pensada,
y la única en que existimos
es la que está dividida
entre la cierta y la errada.*

FERNANDO PESSOA,
“TENGO TANTO SENTIMIENTO” (fragmento)



Inicio este texto con este poema de Pessoa por lo acertado y aplicable que es para el tema de la identidad en las redes sociales, en la era que hoy en día llamamos de la Cuarta Revolución. Mi propósito en este texto es el de denunciar un problema grave que se está presentando por el uso de las redes sociales y por la aparición de algoritmos que nos ofrecen información específica basados en nuestras preferencias pasadas. Este problema tiene que ver con la imagen que nosotros formamos de nosotros mismos y, gracias a ello, con nuestra identidad personal. El problema, dicho en pocas palabras, es el de una forma de autoengaño que se está haciendo omnipresente sin que nos demos cuenta de ello. Para explicar esto, comenzaré por explicar algunas distinciones esenciales del autoconocimiento (o el conocimiento que tenemos de nosotros mismos), para explicar luego dos modelos del autoengaño en la filosofía. Con estas piezas podré, en una tercera parte, construir el argumento que muestra cómo, en este mundo lleno de redes sociales y algoritmos mercantiles, caemos más frecuentemente en el autoengaño de lo que pensamos.

1. Autoconocimiento

Se suele creer que el conocimiento que tenemos de nosotros mismos es un conocimiento al que tenemos un *acceso privilegiado*, es decir, es un conocimiento más certero y menos propenso al error del que tenemos acerca del mundo exterior. Esto se basa en varios hechos comprobables en la vida cotidiana, por ejemplo, al mirar una cierta pintura digo que me gusta la figura de color rojo que se encuentra en la esquina superior derecha del lienzo. Quien me acompaña en el museo puede decir algo acerca de esta figura y es convencional que la llamemos *roja*, pero cuál sea el color que él ve, en su mente, es algo que me es del todo desconocido. Cuando yo pienso en el color rojo un color aparece en mi mente, es a ese color al que llamo *rojo*, pero no sé (ni puedo saber) cuál es el color que aparece en la mente de otra persona cuando piensa en el color *rojo*. ¿Es el mismo tono del mío? ¿Es siquiera similar a lo que yo llamo *rojo*? Al mismo tiempo, parece que hay una cierta *infallibilidad* acerca de mis propios estados mentales, lo cual no ocurre respecto de la información que tengo del mundo exterior. Así, por ejemplo, hay una asimetría entre dos frases aparentemente muy similares pero enormemente diferentes:

- a) Hay un sombrero sobre la mesa.
- b) Creo que hay un sombrero sobre la mesa.

Lo que hace verdadero a (a) es una cierta configuración de los objetos en el mundo. Para que (a) sea verdadero ha de haber cuando menos un sombrero, una mesa y la relación entre los dos debe ser tal que el primero esté posicionado sobre el segundo. Y para que yo sepa esto y pueda afirmarlo mi conocimiento ha de ser inferencial a partir de la evidencia. Así, puedo saberlo porque lo he visto yo mismo o porque alguien me lo ha contado o porque he visto una fotografía tomada hace apenas segundos que me muestra la escena, etc. Todo conocimiento inferencial de este tipo es

falible. Pese a que vi la escena con mis propios ojos, es posible que tenga un grave resfriado y la fiebre me esté haciendo alucinar; quien me ha descrito oralmente la escena podría haberme mentido y las imágenes que he visto podrían ser trucadas y falsas.

Pero lo que hace que (b) sea verdadero no es nada en el mundo, sino algo en mi propia cabeza, en mi mente. Estoy hablando de mis creencias, y (b) será verdadero siempre y cuando yo crea que hay un sombrero sobre la mesa, aún si estoy equivocado acerca de este hecho en el mundo exterior. Y el error parece no presentarse en este caso, pues o bien creo que hay un sombrero sobre la mesa o bien no lo creo, pero parece cuando menos extraño que yo mismo me equivoque acerca de si quiero tal cosa. La razón de ello reside en que este conocimiento que tengo de mis propias creencias no es inferencial, no se basa en evidencia, sino que es directo. Así, aún si estoy alucinando por la fiebre o el alcohol, y creo ver un sombrero sobre la mesa cuando en realidad estoy muy enfermo en la cama de un hospital y no hay allí ni sombrero ni mesa, seguirá siendo verdadero (b), pues es verdad que creo que hay un sombrero sobre la mesa.

Esta asimetría ha sido usada para decir que el conocimiento que tenemos de nosotros mismos es virtualmente infalible y que este acceso privilegiado que tenemos acerca de él hace que sepamos perfectamente todo lo que hay que saber de nuestros propios estados internos (llámense creencias, deseos, anhelos, motivaciones, sentimientos, etc.). Así, pareciera que 1) nadie nos puede conocer mejor de los que nos conocemos nosotros mismos y 2) no podemos estar equivocados acerca de este autoconocimiento.

Sin embargo, esto es algo que ha sido cuestionado desde hace ya varias décadas. El acceso privilegiado que parecemos tener acerca de nuestros propios estados internos no es más que una sensación de *pertenencia* o de *propiedad*, es decir, la sensación de que son *míos* y *de nadie más*, pero qué sea eso que llamo mío, cuál sea su contenido, es algo que, como todo

conocimiento, viene mediado por el lenguaje y, por ende, por una serie de construcciones intersubjetivas que hacen que sea igual de falible que el conocimiento que tenemos del mundo exterior (Hoyos Jaramillo, 2014, cap. IX). Un ejemplo sirve para ilustrar este punto. Hace algunos años me extrañaba el que, recién acabado de comer, sentía hambre. No podía entender cómo, después de un generoso almuerzo, vinera casi de inmediato la sensación de hambre. Una amiga, poco después, me explicó que podía ser gastritis, a lo cual respondí que jamás había tenido, así que no podía saberlo. Ella me explicó, rudimentariamente, cómo se siente la gastritis y lo similar que puede ser a la sensación de hambre, y he de decir que desde entonces puedo identificar con bastante seguridad cuándo tengo hambre y cuándo siento gastritis. Lo importante de esta historia es que estaba equivocado acerca de un estado interno mío. La presunta infalibilidad o inmunidad al error frente a mis propios estados internos era ficticia. Como este puede haber cientos o miles de ejemplos similares. Tomemos el que parece más infalible de todos: el del tono de rojo que veo en mi mente cuando pienso en ese color. Ciertamente parecería que acerca de esto no puedo estar equivocado. Pero cuando ponemos esto bajo la lupa la aparente certeza se desvanece. Digo que veo una mancha roja en mi mente. Pero las palabras *rojo* y *mancha* son palabras cuyo sentido no es estipulado por mí, sino que depende de una aceptación y consenso social. ¿Las estoy usando correctamente? Ahí podría haber un error. Para evitar el error, intento entonces no usar este tipo de palabras, pero entonces ¿cómo me describo a mí mismo lo que veo? Solo puedo decir que veo algo, pero cuando intento describirlo he de usar palabras que consigo traen el potencial error. Para evitar el error he de abandonar el contenido. Pero al abandonar el contenido mi experiencia es vacía (Dancy, 1994, p. 60). Con lo que me quedo es, entonces, con la sensación de que eso que siento, pero que no puedo describir, es mío. Es un sentido de propiedad, de que me pertenece, pero no es más que eso. Cualquier contenido que le añada lo hará falible. Pero sin

contenido no es conocimiento en absoluto, ni siquiera de mí mismo, pues el conocimiento ha de ser *conocimiento de algo*.

Con este tal vez afanoso y posiblemente enredado circunloquio he querido mostrar un punto relativamente sencillo: podemos (y de hecho es común que lo hagamos) equivocarnos acerca de nosotros mismos, esto es, tener un conocimiento fallido de lo que somos, lo que creemos, lo que pensamos y lo que deseamos¹. Esta falibilidad es la que abre la puerta al autoengaño, que será el tópico de la siguiente sección.

2. Autoengaño

En esta sección quiero presentar dos modelos del autoengaño. Uno que llamaré canónico, y que deseo rechazar, y otro que es el que ha sido llamado deflacionista, que es el modelo que suscribo.

El modelo clásico trata al autoengaño como una versión ligeramente modificada del engaño tradicional. En el engaño tradicional tenemos dos sujetos A y B , en donde A desea engañar a B , así que le miente (para mentir, A debe saber cuál es la verdad y cuál es la mentira), con la intención de que B crea la mentira, es decir, con la intención de engañarlo y, por supuesto, para que haya engaño toda esta estratagema ha de ser exitosa (de lo contrario tendríamos la tentativa del engaño, pero no el engaño). Así, en resumen, hay engaño cuando:

1. A cree que P es verdadero.
2. A tiene la intención de engañar a B .
3. A le proporciona a B información acerca de que P es falso.
4. Gracias a la información proporcionada por A , B forma la creencia de que P es falso.

Nótese que en esta situación hay dos cosas dignas de resaltar. En primer lugar, están en juego simultáneamente las creencias de que P es verdadero y

de que P es falso (cada uno de los sujetos tiene una. Y, en segundo lugar, existe una intención de engañar (de A a B). En el modelo de autoengaño canónico la estructura es la misma, solo que ya no tenemos dos sujetos sino uno solo que se engaña a sí mismo:

1. A cree que P es verdadero.
2. A tiene la intención de engañar a A .
3. A le proporciona a A información acerca de que P es falso.
4. Gracias a la información proporcionada por A , A forma la creencia de que P es falso.

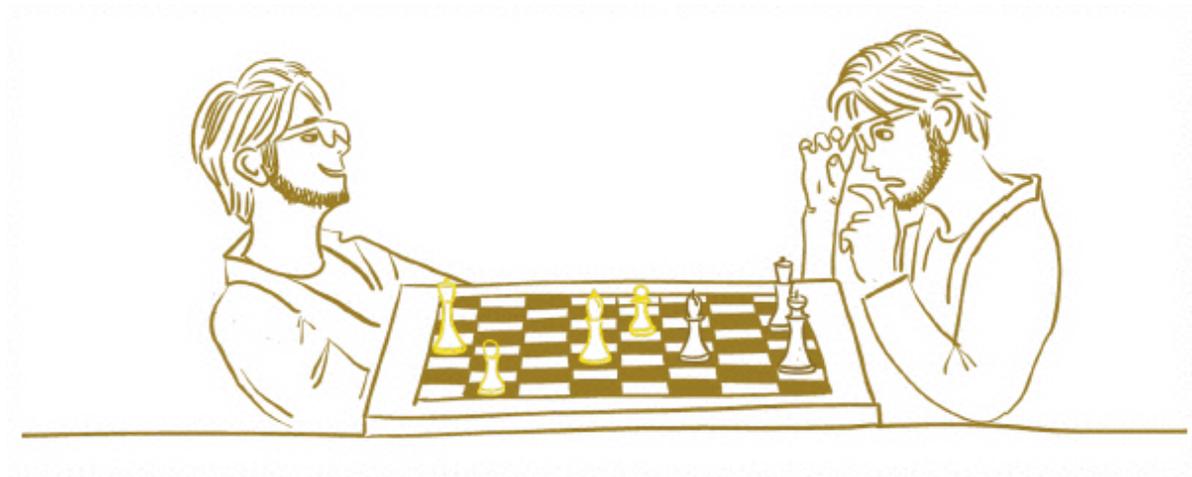
Este modelo conserva las mismas características del modelo de engaño interpersonal que recién vimos. Hay, en primer lugar, creencias opuestas mantenidas simultáneamente (A cree que P es verdadero y el mismo A cree que P es falso) y hay, en segundo lugar, una intención de engañar (A desea engañarse a sí mismo).

He de decir que este modelo de autoengaño, pese a ser el más popular en los estudios filosóficos y psicológicos, me parece del todo inadecuado. Me recuerda a esa maravillosa novela de Stefan Zweig en la que un individuo encarcelado y en confinamiento solitario pasa el tiempo jugando ajedrez mentalmente contra sí mismo, frente a lo cual el mismo Zweig dice:

Si una misma persona juega con las blancas y con las negras, se produce entonces una situación incongruente, en donde un mismo cerebro ha de saber y al mismo tiempo no saber, ha de ser capaz de olvidar completamente cuando juega con las negras, lo que quería y pretendía cuando jugaba con las blancas. Un doble pensamiento como éste presupone en realidad una escisión absoluta de la conciencia, una capacidad de enfocar y desenfocar el cerebro como si fuera un aparato mecánico; querer jugar contra uno mismo representa, en definitiva, una paradoja tan grande en ajedrez como querer saltar sobre la propia sombra. (Zweig, 2003, p. 65-66)

Ciertamente este modelo del autoengaño presenta paradojas. ¿Cómo puede alguien creer que algo es verdadero y que es falso? ¿Cómo puede

alguien tener la intención de engañarse con información que sabe que es falsa? No han sido pocos quienes han intentado defender este modelo. Amélie Okseberg Rorty, por ejemplo, ha propuesto un modelo de la mente humana encapsulado o fragmentado, en donde la información que está en una de sus cápsulas puede pasar desapercibida a la conciencia cuando examina otra de las cápsulas (permitiendo así que en una misma mente haya la creencia de que P es verdadero —en una cápsula— y de que P es falso —en otra—). Como la propuesta de Okseberg Rorty hay muchas otras que proponen funcionamientos cerebrales especiales, por ejemplo, Robert Audi (1988) propone que una de las creencias opuestas ha de ser *inconsciente*. No sé bien qué significa que haya creencias inconscientes, pero ciertamente se nota que es un esfuerzo grande por acomodar el cerebro al modelo de autoengaño (para más modelos similares cf. McLaughlin y Rorty, 1988).



No es necesario entrar en una larga e intrincada refutación de estas propuestas; basta simplemente con proponer un modelo del autoengaño que no sea tan demandante, que no implique estas enredadas paradojas y que no requiera de estas peripecias acerca de los modelos mentales para que el autoengaño tenga alguna plausibilidad. Afortunadamente tal modelo existe y es el modelo deflacionista de Alfred R. Mele.

Este modelo pretende explicar la misma fenomenología del autoengaño, pero de una manera más simple, en dos sentidos importantes: 1) no requiere que haya dos creencias opuestas simultáneamente sostenidas por el mismo individuo y 2) no requiere que haya una intención de engañarse a sí mismo. Es decir, los dos puntos más problemáticos del modelo canónico están ausentes de este modelo deflacionista, lo cual lo hará, sin duda, más aceptable *prima facie*.

La idea de Mele es muy ingeniosa. Lo que solemos llamar casos de autoengaño son casos en los que un sujeto cree cierta cosa, cuando toda la evidencia muestra que eso que él cree es falso. Ejemplos abundan: el marido que cree que su esposa le es fiel pese a que hay evidencia, por toneladas, que le muestra que no lo es. O la persona que cree que es generosa, pese a que toda la evidencia que posee apunta a que no lo es, ni con su tiempo ni con su dinero. Para Mele estos casos se presentan porque los sujetos hacen un tratamiento sesgado de la evidencia que tienen a su disposición. Y este tratamiento es sesgado porque hay una motivación, usualmente emotiva, que los lleva, sin que se den cuenta, a ello. Así, ellos hacen un tratamiento incorrecto de la evidencia, gracias a la motivación que tienen, y extraen de allí una creencia falsa, mientras que las personas que no tienen esa motivación y examinan la misma evidencia, extraen de allí la creencia correcta.

Un ejemplo puede ayudarnos a explicar este punto. Julián está enamorado perdidamente de Susana. La busca por todas partes, le escribe mensaje y le propone que se vean en algún lugar solo. Susana no responde sus mensajes, le es evasiva y, cuando tiene que verse con él, se asegura de estar siempre acompañada por un grupo grande de amigos para no estar a solas con Julián. Cualquiera de nosotros, que no tiene sentimiento romántico alguno por Susana, extraerá de esta evidencia que ella no quiere tener ninguna relación romántica con Julián. Sin embargo, él tiene algo que nosotros no tenemos. Su enamoramiento por Susana lo lleva a examinar

sesgadamente la evidencia. Esta motivación hace que, cuando él observa la evidencia, mire solo las partes de ella que concuerdan con su motivación (con su enamoramiento). Así, de la misma evidencia de la cual nosotros extraemos la creencia de que Susana no desea una relación romántica con él, él extrae la creencia de que Susana está loca por él. Él ve la evidencia como mostrando que ella está tan perdidamente enamorada de él que no se atreve a responder sus mensajes por miedo a quedar en evidencia. Que no confía en ella misma para contenerse si está a solas con él y, por ello, invita a otras personas cada vez que tienen que verse para no sucumbir a sus deseos, etc. Es claro cómo, ante la misma evidencia que nosotros, la presencia de ese factor motivacional ha hecho que Julián haga un tratamiento incorrecto y sesgado de la evidencia y extraiga de ella una creencia falsa².

Ha de notarse algo importante en este ejemplo: en ningún momento se ha dicho que Julián desea engañarse, ni se ha dicho que Julián cree que Susana no desea una relación con él y que cree, simultáneamente, que Susana está perdidamente enamorada de él. En este modelo deflacionista no hay, como habíamos dicho, ni la intención de engañarse, ni creencias opuestas sostenidas simultáneamente. Julián no forma una creencia falsa porque desee engañarse. Él no sabe que Susana no desea una relación con él, simplemente ha tratado mal la evidencia debido a sus sentimientos por Susana y de ahí ha extraído una creencia incorrecta.

Como se puede ver, el modelo es bastante sencillo. Su estructura sería la siguiente:

1. *A* adquiere una creencia acerca de *P* que es falsa.
2. *A* trata los datos relevantes, o al menos aparentemente relevantes, para el valor de verdad de *P* de una manera motivacionalmente sesgada.
3. Este tratamiento sesgado es una causa no desviada de que *A* adquiera

la creencia falsa de que P .

4. Los datos a disposición de A en ese momento proveen mayor evidencia para una P verdadera que para una P falsa (Mele, 2009).

¿Podemos llamar a esto *autoengaño*? No veo razón alguna para decir que no. El modelo canónico del autoengaño (el anterior que hemos visto) ha sido creado justamente para explicar casos en los que, pese a la abundante y contundente evidencia disponible, un sujeto forma una creencia opuesta a ella. Este modelo deflacionista puede explicar exactamente esos mismos casos, de una manera menos demandante porque no incluye las paradojas del engaño de sí mismo y de las creencias encontradas. Lo que importa es el fenómeno (los casos), que permanece igual, y lo que hemos cambiado es el modelo mediante el cual damos cuenta de él o lo explicamos. Estamos explicando lo mismo, pero con un modelo diferente. Si a estos casos se los llamaba antes *autoengaño* no hay razón para no usar este mismo apelativo, aún si la intención de engañar no está presente en el segundo modelo.

Hasta el momento el recorrido por este texto nos ha llevado por una sección acerca del autoconocimiento en la cual he defendido que, pese a lo que usualmente se cree, podemos (y de hecho lo hacemos con frecuencia) fallar y equivocarnos acerca de lo que creemos, pensamos, deseamos y sentimos. Esto abre la puerta a la posibilidad del autoengaño, lo cual nos ha llevado por una segunda sección en la que hemos explorado dos modelos propuestos para explicar este fenómeno. He rechazado el primero (el canónico) y he suscrito el segundo (el deflacionista). En la sección que sigue utilizaré este modelo para mostrar cómo podemos tener problemas de autoengaño con nuestra identidad personal.

3. Identidad

Hay varios tipos de identidad. La identidad material, por ejemplo, es la que adjudicamos a una cosa que vemos en dos momentos diferentes. Cuando

decimos, por ejemplo, de una mesa a la que vemos por la mañana y luego la volvemos a ver en la tarde que es *la misma*, lo que estamos haciendo es afirmar su identidad en el tiempo, esto es, su permanencia como un mismo ente. El tratamiento sistemático más conocido de la identidad en la filosofía moderna se debe a John Locke quien, preocupado por el tema de cuándo podemos afirmar la identidad de algo, estableció tres tipos de identidad: “Una cosa es ser la misma substancia, otra ser el mismo hombre, y una tercera ser la misma persona. Si persona, hombre y substancia significan tres ideas diferentes, en tanto que la idea pertenece a un nombre, entonces así tendrá que serlo también la identidad” (Locke, 2000, p. 315). La identidad de las substancias (entiéndase objetos inanimados) la da el hecho de que sus partes son las mismas y permanecen en la misma configuración. Así, la mesa que vemos en la mañana en nuestro comedor y la mesa que vemos en la tarde allí mismo son la misma, porque vemos la misma cantidad de materia, con las mismas piezas y en el mismo orden. Si la mesa de cuatro patas en la mañana, por la tarde tuviera solo una central o si por la mañana fuera de madera, pero por la tarde fuera de vidrio, no diríamos que es la misma. Algo diferente ocurre con los seres vivos. Al plantar una semilla de roble, verla germinar a los pocos días y, años después, ver el gran roble en el que se ha convertido, decimos que se ha tratado del mismo roble, es decir, afirmamos una identidad. Pero ciertamente no aplica aquí el criterio de la identidad de substancias, pues al crecer el roble tiene más materia, más piezas y ciertamente estas están en un orden distinto que cuando era una semilla, cuando era un retoño y cuando ya es un árbol adulto. La identidad la afirmamos, según Locke, porque vemos una sucesión de materia y de configuraciones que están todas asociadas a una misma vida. Es la permanencia de la vida la que nos permite afirmar la identidad de esa planta que tanto cambia: “lo que constituye la unidad de la planta es la organización de sus partes en un cuerpo coherente que participa en una vida común, una planta continúa siendo la misma mientras siga

participando de la misma vida”. (Locke, 2000, p. 313). Y esto que vale para las plantas, aplica también para los animales y, en tanto que los seres humanos somos también animales, aplica para nosotros *en tanto que animales*. Es a esto a lo que Locke llama la identidad *como hombre*. Pasamos de ser un feto, a ser un bebé, un niño, un adolescente, un adulto y luego un anciano, pero todos estos cambios están asociados a una misma vida y, por ello, podemos decir que hemos sido siempre el mismo ser humano.



Ahora bien, Locke distingue entre la identidad del ser humano y la identidad de la persona. No son lo mismo. Los seres humanos tenemos, además de nuestro cuerpo como organismo vivo, una vida mental³. Es muy común que digamos cosas como *él ya no es el mismo que era antes*, cuando nos encontramos con un amigo que no veíamos después de mucho tiempo. Ciertamente el amigo, como cuerpo vivo, ha tenido una misma vida y, entonces, es el mismo ser humano. De manera que cuando decimos que ya no es el mismo, cuando negamos su identidad, lo que queremos decir es algo diferente. Lo que queremos decir es que ya no es *la misma persona*.

Nuestra vida mental es la que constituye el tipo de persona que somos. Nuestras creencias, nuestros deseos, nuestras intenciones y las acciones que de ellas se derivan, entre otras cosas. Es porque la identidad como ser humano y la identidad como persona son diferentes, tanto que podemos pensar en la posibilidad de la transferencia mental (*mind uploading*), esto es, la posibilidad de que algún día podamos transferir nuestra vida mental a un computador cuyas partes pueden irse reemplazando y evitar así las limitaciones de tiempo que nos da el inevitable decaimiento de nuestro cuerpo. De lograrse esto, podríamos vivir casi que eternamente. Seríamos, en ese caso, la misma persona, pero no seríamos ya el mismo ser humano (ya no tendríamos una vida biológica y un cuerpo biológico asociado a ella). Regresando a Locke, al preguntarse qué es lo que nos permite decir de una persona que es la misma, encuentra esta solución:

Un ser pensante, inteligente, dotado de razón y reflexión, y que puede considerarse a sí mismo como el mismo, como una misma cosa pensante en diferentes tiempos y lugares, lo hace tan sólo en virtud de tener conciencia, que es algo inseparable del pensamiento y que, me parece, le es esencial [...] Y hasta el punto que ese tener conciencia pueda alargarse hacia atrás para comprender cualquier acción o cualquier pensamiento pasados, hasta ese punto alcanza la identidad de la persona: es el mismo *sí mismo* ahora que era entonces. (Locke, 2000, p. 318)

La conciencia y la memoria son las que nos dan, entonces, nuestra identidad personal. Soy la misma persona porque 1) soy consciente de mis deseos, de mis creencias, de mis anhelos y experiencias y 2) porque puedo recordar esta conciencia que he tenido en momentos pasados. Así, si bien ya no me gusta un cierto caramelo que me encantaba en mi infancia, recuerdo bien que me gustaba y por ello puedo decir que hoy soy la misma persona a la que antes le gustaba ese caramelo⁴. Los controvertidos casos de *personalidad múltiple* (sea que existan realmente o no) solo serían posibles si las dos personalidades (o las dos ‘personas’ que habitan el mismo cuerpo) no comparten una conciencia y una memoria. Pues, en el momento en que

una personalidad pueda acceder a la memoria de la otra, recordaría las acciones como *suyas*, y así colapsarían las dos en una misma personalidad. Se diluiría la escisión y quedaría solo una.

Ahora bien, teniendo claro en qué consiste la identidad personal, podemos explorar un poco más la complejidad que rodea a este tipo de identidad. Ojalá las cosas fueran tan sencillas como Locke las presenta, pero la verdad es que la vida mental de los seres humanos (y con ellos su identidad personal) es en extremo compleja. Una parte esencial de esta complejidad es que cuando hablamos de quiénes somos, del tipo de identidad personal que tenemos, podemos hablar de dos cosas: la primera de ellas es quiénes somos realmente; la segunda, quiénes creemos que somos. Recordemos el poema de Pessoa con el que iniciamos este capítulo. Allí se habla de una doble vida, eso es justamente a lo que deseo hacer referencia aquí. Luciano Floridi describe esta distinción de la siguiente manera:

En la filosofía de la mente hay una distinción bien establecida entre quiénes somos [...] y quiénes pensamos que somos [...]. Sobra decir que hay una diferencia crucial entre ser Napoleón y creer ser Napoleón. Estos dos *yoes* florecen sólo si se apoyan mutuamente en una relación saludable. No solo quienes creemos que somos debe ser algo cercano y basado en quienes realmente somos, sino que quienes realmente somos es algo maleable e influenciado por quienes creemos que somos o quisiéramos ser. Si usted cree que es una persona segura de sí misma, es muy probable que se convierta en una persona segura de sí. (2014, p. 60)⁵

Quiero llamar *yo real* a quien somos verdaderamente, y *yo ideal* a quien creemos que somos. Creo que no es difícil de aceptar que tenemos dos concepciones de nuestro propio yo. Justamente porque podemos equivocarnos acerca de quiénes somos (*cf.* § 8.1) nos es común sorprendernos a nosotros mismos cuando, por ejemplo, descubrimos que somos celosos (pese a que creíamos que no lo éramos). En un caso nuestro *yo ideal* era el de no ser celoso, pero gracias a alguna evidencia contundente

nos damos cuenta de que en realidad sí lo somos, es decir, el *yo real* diferiría del *yo ideal*. Como bien lo dice Floridi, una identidad sana es aquella en donde no hay una gran distancia o brecha entre los dos *yoes*. Siempre creemos que somos un poco más juiciosos de lo que en verdad somos, o tal vez que somos más pacientes y comprensivos de lo que en verdad somos. Pero mientras las distancias entre el *yo real* y el *yo ideal* no sean muy grandes, no hay problema alguno; incluso, como también dice Floridi al final de la cita, en ocasiones esta mínima diferencia es la que permite que mejoremos, que hagamos un esfuerzo y terminemos por convertirnos en más juicioso, más aplicados o más comprensivos.

Hasta el momento hemos logrado mostrar que, ya que podemos equivocarnos acerca de nosotros mismos (el conocimiento que tenemos de nosotros, o autoconocimiento, puede ser errado), es perfectamente comprensible (y hasta común) que nos equivoquemos acerca del tipo de personas que somos. Quienes creemos que somos, a veces, termina por no ser quienes realmente somos. En la siguiente sección procuraré mostrar cómo esta brecha entre el *yo ideal* y el *yo real* puede dar pie al autoengaño, para mostrar, luego, cómo en nuestra vida en redes sociales el peligro del autoengaño se multiplica exponencialmente.

4. Las redes sociales al servicio del autoengaño

En la sección anterior hemos visto que, cuando hay una brecha o distancia muy grande entre el *yo ideal* y el *yo real*, el sujeto puede ser propenso al autoengaño. Intentaré explicar ahora por qué ocurre esto. Lo primero que hay que tener en cuenta es que mantener o abandonar una creencia puede tener un costo. Si, por ejemplo, estoy en una comunidad en donde ser un hombre rudo y no demostrar jamás miedo es un ideal de aceptación, me será mucho más fácil ser aceptado, tener relaciones sociales y, en general, sentir que pertenezco a un grupo si proyecto esa imagen. Esa podría convertirse en la imagen de mi *yo ideal*. Si, por alguna razón, la distancia

entre ese *yo ideal* y el *yo real* es muy grande (supongamos que mi *yo real* es el de una persona temerosa, poco aventurera, llena de fobias, etc.), fijar esta imagen como verdadera, es decir, aceptar que mi *yo ideal* no es en realidad lo que me representa será muy costoso. Con él no solo se viene una creencia que tengo acerca de mí mismo, una creencia que me agrada, sino que además podrían venirse abajo mis relaciones sociales, puedo ser excluido de un grupo.

Muchas veces nos identificamos con nuestro *yo ideal*, es decir, creemos que nuestro *yo ideal* es quienes somos. Y aceptar que no somos como creemos que somos es costoso, es descubrir que hemos creído erradamente algo, que creíamos ser una persona que, en realidad, no concuerda con nuestra realidad. Es común que esto ocurra (claro está, en pequeñas cosas). Por ejemplo, como académicos solemos creer que somos más inteligentes y brillantes de lo que somos. Solemos decirnos que, si nos dedicáramos en serio a trabajar en un artículo o libro, sin duda sería brillante, pero que preferimos balancear el tiempo entre la familia, el trabajo, etc. Es decir, creemos que tenemos capacidades enormes cognitivas y académicas. En parte ello se debe a que la academia es un ambiente competitivo, en donde todo el tiempo hay disputas y mediciones y ser el mejor es algo muy importante. Por ello, cuando en realidad dedicamos enormes esfuerzos a un artículo o un libro, dejando todo lo demás de lado, y un evaluador lo rechaza, solemos pensar de inmediato que el evaluador no ha leído con atención el texto, que no ha entendido la propuesta, que no está capacitado por evaluar el texto o alguna otra excusa con la cual ajustamos la evidencia abrumadora de que nuestro artículo no es tan bueno y la hacemos parecer como si fuera culpa del evaluador. Con esto protegemos una imagen que es muy importante para nosotros, es una imagen que nos permite continuar en la academia sin sentirnos como un fraude. Aceptar la realidad puede ser muy costoso, muy doloroso.

Y en este ejemplo hemos visto cómo funciona el modelo deflacionista a

la perfección. Tengo un deseo, a saber, ser aceptado en el mundo académico, que es altamente competitivo. Este deseo de pertenecer a la academia es lo que hace que, ante la evidencia que muestra que no soy tan bueno y que (tal vez) no pertenezco allí, haga un tratamiento sesgado y extraiga una creencia que otra persona, sin este mismo deseo, no extraería (extraería la contraria). Así, mientras cualquier otro extraería “no eres tan bueno como crees”, yo extraigo “soy excelente pero el evaluador es muy malo y no ha leído bien mi escrito”.

Es importante resaltar una cosa aquí: el sujeto que tiene este deseo *deja de ser un buen evaluador de la evidencia*. Parece redundante que lo diga, pues hace parte del modelo deflacionista, pero es necesario, porque debemos hacer la pregunta de qué es lo que nos saca de este estado de autoengaño. Y mi respuesta es: no son los hechos, son las otras personas. Los hechos (entendiendo por esto situaciones inertes) no se resisten a una mala interpretación. Una hoja de papel en la que se dice que mi texto es rechazado y no es bueno no se resiste a que yo la vea como conteniendo información injusta o falsa. Puedo decir que ese papel dice basura y el papel no responderá nada. Lo mismo ocurre con cualquier otro hecho. Pero las personas sí se resisten. Si yo digo que soy el ser más generoso del planeta (y no lo soy), mis amigos, mis familiares, mis seres cercanos me enfrentarán, dirán que no es así y me recordarán situaciones en las que no he sido, para nada, generoso con mi tiempo, con mi dinero, etc. Yo puedo equivocarme e interpretar los hechos (esto es, tratar la evidencia) de una manera incorrecta. Pero los hechos siempre estarán allí, inermes. Pero no ocurre lo mismo con las personas. Aquellas que se interesan por mí me enfrentarán, porque no quieren que viva una mentira. Me mostrarán la evidencia y me llamarán para que la vea bajo una nueva luz.

La única manera de escapar de esta resistencia que me hacen otras personas es cortando las relaciones con ellas. Este corte no debe ser total, de hecho, es muy común que sea apenas parcial, pero el punto es que

rompemos una cierta relación, el menos en un tema específico, con la persona que se resiste. Así, si mis opiniones acerca de política son erradas, porque he hecho una evaluación sesgada de la evidencia gracias a algún deseo particular que tengo, y una persona me recrimina que son erradas estas creencias, y se resiste a mis juicios, es muy probable que yo simplemente diga *con fulano no se puede hablar de política*, y así lo excluyo de estas conversaciones, protegiendo además mis creencias, mis deseos y la imagen que tengo de mí mismo. Seguramente puedo seguir hablando con él de fútbol, de literatura y otros temas, pero corto mi relación en lo que tiene que ver con política.

Creo que el autoengaño es mucho más común de lo que creemos y creo que todos, de una manera u otra, tenemos un comportamiento basado en el autoengaño similar a lo que he descrito aquí. Eso no es, en principio, grave. Se convierte en grave cuando, como he dicho, la brecha entre el *yo ideal* y el *yo real* es muy grande. En esos casos la resistencia de quienes me rodean ha de ser más grande. Supongamos que tengo un problema con la bebida, pero no lo veo porque tengo algún deseo que me hace evaluar sesgadamente la abrumadora evidencia de que tengo este problema. Supongamos que los problemas con el alcohol hacen que sea una persona agresiva cuando toma (que es casi siempre) y hasta violenta con su familia. Si mis amigos, mi familia y mis seres cercanos me enfrentan todo el tiempo con este problema, mi única escapatoria es cortar las relaciones con ellos. Pero este corte de relaciones ha de ser mucho más severo que simplemente no hablar de un tema. Podemos asumir que la conversación sobre mis problemas con el alcohol surge cada vez que tomo, y que esto es muy, muy frecuente, así que para evitar la conversación tendré que cortar muchas relaciones con mis seres cercanos. Poco a poco me voy aislando. Y aquí solo hay dos desenlaces posibles. Uno, que es el más trágico, es que me aísló definitivamente y dejo que la bebida me consuma. Ensimismado en la historia o narrativa que he creado según la cual no tengo un problema. O

bien, que es lo más común, mi necesidad de ser un ser social hace que extrañe estas relaciones, y es esta manera de extrañar a los otros lo que hace que mire la evidencia de manera diferente y note que tengo un problema.

Aquí hay dos elementos importantes. Por un lado, el deseo que tengo (sea el que sea), que me lleva a examinar de manera errada la evidencia, se enfrenta ahora a otro deseo (el de la compañía de mis seres queridos). Y, por el otro, el costo que tenía antes aceptar mi *yo real* se compara ahora con un costo potencialmente mayor, a saber, el de perder mis relaciones sociales, alejarme de la gente que quiero. En estas circunstancias puede ser más costoso perder las relaciones sociales y, por ello, reevaluamos la evidencia y notamos, por fin, el problema. Como decía antes, no es tanto la evidencia misma, la verdad o los hechos, lo que nos puede sacar de la situación de autoengaño, sino los otros, las personas que están a nuestro lado, su cariño, las relaciones con ellos, eso es lo que ayuda a que veamos las cosas de manera diferente.

He mostrado hasta aquí cómo la división entre dos *yoes* nos hace propensos al autoengaño, también he intentado mostrar cómo no es *la verdad*, o *los hechos* lo que nos permite salir de allí, sino los otros, nuestros amigos, nuestros seres cercanos. En lo que resta de este texto intentaré mostrar por qué las redes sociales en la Revolución 4.0 aumentan exponencialmente el riesgo de hacer en autoengaño y, a la vez, hacen aún más difícil salir de esta situación.

Son tres los elementos de estas redes que contribuyen a este fenómeno: 1) el que en ellas los amigos sean virtuales, 2) el que en ellas podemos mostrar un *yo* filtrado y 3) los algoritmos con los cuales las compañías de servicios en internet intentan vendernos productos.

Vayamos al primer elemento. Los amigos en redes sociales pueden ser de dos formas. O bien se mantiene un contacto virtual con un amigo con quien se ha tenido una amistad presencial, o bien la amistad es enteramente virtual. El segundo grupo es mucho más numeroso que el primero. Tenemos

amigos que hacemos jugando juegos en línea y a quienes jamás en la vida conocemos. Gente que nos pide ser amigos en Facebook y con quienes compartimos nuestras imágenes, eventos, fotos de comida y mascotas, comentamos cosas, jugamos, pero jamás nos conocemos personalmente. Cuando vemos a *youtubers* notamos que tienen desde mil hasta cinco o seis millones de seguidores, con quienes comparten sus vidas, para ellos hacen sus videos, cuentan historias, pero a quienes rara vez conocen.

El segundo elemento es que la imagen de nosotros mismos que ofrecemos *online* es una imagen parcial. En redes sociales solemos compartir y hacer públicos nuestros logros, nuestros triunfos, pero rara vez mostramos nuestro lado vulnerable, nuestros errores, nuestros remordimientos. Creamos una suerte de personaje y filtramos lo que le mostramos al mundo de él. Puedo desconectarme de la red social mientras me calmo o mientras pienso algo y volver a conectarme luego. Si quiero parecer inteligente puedo consultar una enciclopedia o cientos de páginas web para exponer un dato curioso y rebuscado. Nada de esto puedo hacerlo con mis amigos presenciales. Ellos ven cómo reacciono porque rara vez puedo esconderme. En la excursión del colegio todos vimos cómo manejamos la frustración, vimos la timidez al acercarnos a las jovencitas del otro colegio que estaba también en excursión en el mismo sitio, etc. Claro, podemos intentar filtrar mucho de lo que mostramos, pero es mucho más difícil y, tarde o temprano, tendrán que encontrarse con esa versión no filtrada de nosotros mismos.

Finalmente, es ya común el conocimiento de que las grandes empresas y, con ellas, las plataformas que albergan a las redes sociales utilizan algoritmos diseñados para almacenar grandes cantidades de información de los usuarios y, a partir de esta, ofrecerles productos que les interesen. Se trata de algoritmos complejos que se basan en lo que los usuarios han visto, por ejemplo en Amazon, y lo que han comprado, para recomendarles otros productos que podrían gustarles (Linden *et al.*, 2003). Mucho más famoso

es el algoritmo con el que Netflix analiza la información que acumula de sus usuarios para recomendarles contenido (Gómez-Uribe y Hunt, 2015), pero estos algoritmos se pueden encontrar hoy en día en un sinfín de plataformas (Gertz, 2018). Cada vez que nosotros, como usuarios, realizamos alguna acción en la red, esta queda registrada. Netflix, por ejemplo, almacena todos los movimientos que hacemos en el menú. Cuánto tiempo miramos la sinopsis de una película, si vemos series entre semana y películas los fines de semana, cuándo abandonamos una serie o una película, qué géneros vemos en el día y qué géneros en la noche, etc. Un artículo de la revista *Wired* de agosto de 2017 (Plummer, 2017) calculaba que había 100 millones de usuarios de Netflix en esa fecha. Asumiendo, como es más que razonable, que una misma cuenta la usan varias personas, para ese entonces Netflix contaba con unos 250 millones de usuarios con perfil activo. Imaginemos ahora a cada uno de estos usuarios mirando el menú, iniciando una serie que abandona a los pocos minutos, mirando la sinopsis de una película, y cada uno de estos movimientos siendo registrado para luego ser analizado. Las cantidades de datos son abrumadoras y el algoritmo las utiliza para buscar lo que más nos gusta y ofrecernos cosas que nos gusten también, es un algoritmo mercantil enteramente, la idea es vender más, pero ahora de manera más dirigida y personalizada. Lo mismo ocurre en Amazon.com y lo mismo ocurre en Facebook.

Con estos tres elementos tenemos una receta de cuidado para el autoengaño. En primer lugar, ya que podemos ofrecer un yo filtrado a las redes sociales, la distancia entre el *yo ideal* y el *yo real* es muy fácil que se agrande más y más. Las comunidades en internet no son, en algunas características, muy diferentes de las comunidades sociales personales. Hay un deseo de pertenecer a ellas, de ser aceptado, de sobresalir, etc., tal y como ocurre con los grupos sociales del colegio, del trabajo y de la universidad. La diferencia estriba en que a estos grupos podemos entrar con una suerte de máscara que podemos ir alimentando con el tiempo. Y no nos

ayuda el hecho de que los algoritmos estén diseñados para ofrecernos cosas que nos gusten, en vez de cosas que nos reten y nos enfrenten. Regresemos por un momento al ejemplo que vimos más arriba de las discusiones políticas. Si yo he hecho una mala evaluación de la evidencia, debido a algún deseo que tengo, no ayuda a que corrija mi error el que haya algoritmos en Facebook que busquen mostrarme únicamente posiciones similares a la mía. El algoritmo es mercantil, busca lo que me gusta y me ofrece más de lo mismo. No está diseñado para enfrentarme a posiciones divergentes de la mía. Es por eso que en discusiones como la del Brexit, la del referendo acerca de los diálogos de La Habana y las elecciones presidenciales en los Estados Unidos, a la gente que tenía una marcada posición acerca de uno de los bandos solo le aparecían noticias que corroboraban su posición y no las que la controvertían. Esto hace que quien está falsamente convencido de algo tenga más evidencia que soporta su convencimiento, en vez de exponerse a la evidencia que muestra que está equivocado. Es como si las redes sociales en sí mismas filtraran la evidencia para que nos afiancemos en nuestro error.

Las noticias y la información en las redes sociales se comportan igual que los hechos en el mundo fuera de la red. No se resisten a la malinterpretación. Pero estaríamos medianamente a salvo si las personas se nos resistieran y, tal como ocurre en los casos fuera de la red, fueran ellas quienes nos ayudaran a salir de nuestro estado de autoengaño. Sin embargo, hay algo que podemos hacer en el mundo virtual, que no podemos hacer en el físico: podemos bloquear personas. Si en redes sociales alguien me increpa por algo que he dicho, puedo bloquear a ese usuario, de la red, de mi cuenta, de mi correo electrónico, etc. En el mundo físico esto no es tan fácil. Si quien me increpa es alguien de mi trabajo, por mucho que no desee verlo me lo encontraré en la oficina, en el almuerzo, en el retiro de la empresa. Si es mi primo, lo veré en las reuniones familiares, si son mis amigos, los veré en reuniones con otros amigos que tenemos en común, etc.

En el mundo físico no es tan fácil ‘bloquear’ a alguien. En las redes, en cambio, con un clic podemos hacer desaparecer a una persona por completo.

Así, los tres elementos se confabulan para hacer del peligro del autoengaño algo mucho más fuerte. 1) Puedo presentar una imagen de mí que diste mucho del yo real; 2) los algoritmos que me presentan información, comentarios y noticias van a afianzar mis creencias en vez de controvertirlas y ponerlas a prueba y 3) las personas que podrían ejercer una resistencia son fácilmente bloqueables. Y, a diferencia del mundo físico, entablar relaciones con otras personas es mucho más fácil. No se corre el riesgo, o al menos no da la impresión de ser un riesgo apremiante, de aislarse, pues es fácil conseguir seguidores, ‘amigos’ de Facebook y personas con quienes jugar un juego *online*. En mi caso particular, he de decir que amigos, amigos de verdad, he de tener unos diez (y me siento afortunado por ello), pero tengo cerca de 300 en Facebook (que es poco para estándares contemporáneos). Las amistades del mundo físico me han costado mucho tiempo, para entrar en confianza, construirlas y afianzarlas. Algunas de ellas son amistades de más de veinte años. En cambio, los 300 amigos de Facebook los tengo desde hace poco más de dos años que creé mi cuenta.

No es mi deseo comparar los amigos presenciales con los virtuales, ni demonizar las redes sociales o siquiera sugerir que las amistades virtuales son menos valiosas que las presenciales. Nada de ello. Hay sin duda amistades virtuales tan valiosas como las presenciales, e incluso en ocasiones mucho más. Mi punto es mucho menos polémico, simplemente he querido mostrar cómo funciona el autoengaño y visibilizar que, en las redes sociales de la Cuarta Revolución, los elementos que llevan al autoengaño se encuentran exponencialmente fortalecidos, a la vez que aquellos que nos permiten salir de esta situación engañosa están sumamente debilitados. Hay, pues, un riesgo inminente frente al cual hay que estar

alerta.

Referencias

- Audi, R. (1988). Self-Deception, Rationalization, and Reasons for Acting”. En B. P. McLaughlin y A. O. Rorty (Eds.). *Perspectives on Self-Deception* (pp. 92-120). University of California Press.
- Cassam, Q. (Ed.). (1994). *Self-knowledge*. Oxford: Oxford University Press.
- Dancy, J. (1994). *Introduction to contemporary epistemology*. New Jersey: B. Blackwell.
- Floridi, L. (2014). *The 4th revolution: how the infosphere is reshaping human reality* (1.^a ed.). Oxford: Oxford University Press.
- Gertler, B. (2010). *Self-Knowledge*. Nueva York: Routledge.
- Gertz, N. (2018). *Nihilism and technology*. Londres: Rowman & Littlefield International.
- Gómez-Uribe, C. A. y Hunt, N. (2015, diciembre). The Netflix Recommender System: Algorithms, Business Value, and Innovation”. *ACM Transactions on Management Information Systems*, 6(4), 1-19.
- Hoyos Jaramillo, L. E. (Ed.). (2014). *Ensayos de filosofía práctica y de la acción* (1.^a ed.). Bogotá: Universidad Nacional de Colombia.
- Linden, G. *et al.* (2003, enero). Amazon.Com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 7(1), 76-80.
- Locke, J. (2000). *Ensayo sobre el entendimiento humano*. México: Fondo de Cultura Económica.
- McLaughlin, B. P. y Rorty, A. (Ed.). (1988). *Perspectives on self-deception*. Los Ángeles: University of California Press.
- Mele, A. R. (2009). Have I Unmasked Self-Deception or Am I Self-Deceived? En C. W. Martin (Ed.), *The Philosophy of Deception*. Oxford: Oxford University Press.
- Patarroyo, C. G. (2009, enero). Correspondencia entre Locke y Molyneux acerca de la identidad personal y el derecho a castigar justamente a un

ebrio que no es consciente de sus acciones. *Ideas y Valores*, 58(139), 145-59.

Plummer, L. (2017, agosto). This is how Netflix's top-secret recommendation system works. *Wired UK*. Recuperado de <https://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like>

Strandberg, H. (2015). *Self-Knowledge and Self-Deception*. Houndmills: Palgrave Macmillan.

Zweig, S. (2003). *Novela de ajedrez*. Barcelona: Acantilado.

Notas

¹ Esta ha sido, sin duda, una exposición muy somera y rápida de un tema en extremo complejo como es el autoconocimiento. Para un tratamiento mucho más profundo y comprensivo son muy recomendables los libros de Cassam (1994) y Gertler (2010).

² El tipo de tratamiento sesgado de la evidencia que ha hecho Julián se conoce como *malinterpretación positiva* y está ampliamente estudiado por la psicología y la filosofía aplicada. Otros tipos de tratamiento sesgado de la evidencia son: *malinterpretación negativa*, *atención selectiva* y *recopilación selectiva de evidencia* (véase, Mele, 2009).

³ No deseo suscribir aquí ningún tipo de dualismo, es decir, no deseo afirmar, como lo hacía Descartes, que la mente es un tipo de sustancia diferente e independiente del cuerpo. Simplemente deseo decir que la identidad *animal* que tenemos como seres humanos deja por fuera un aspecto importante de nuestra vida, que es la vida mental. Y esto es verdadero sea uno monista o dualista.

⁴ No deseo afirmar que la propuesta de Locke carezca de problemas. Pero lo esencial de ella (que el tipo de *personas* que somos depende de nuestra vida mental y no de nuestro cuerpo) es lo que me interesa rescatar y ante esto no hay mayor controversia. Sin embargo, si el lector desea conocer una de las controversias más famosas al respecto, puede leer la correspondencia que Locke sostuvo con su colega Will Molyneux acerca de cómo se puede afirmar la identidad de una persona que, estando ebria, comete un crimen y luego, por la ingesta de alcohol, cuando está sobria no recuerda en absoluto haber cometido el crimen (véase, Patarroyo, 2009).

⁵ De manera muy similar, Strandberg habla de un “yo verdadero” y un “yo fabricado”. Puede verse su explicación de cada uno en el capítulo 5 de su libro *Self-Knowledge and Self-Deception* (2015).