

El auto-conocimiento como expresión

Ángel GARCÍA RODRÍGUEZ

Universidad de Murcia

1. Introducción

En la filosofía contemporánea hay un debate abierto acerca del auto-conocimiento; es decir, el conocimiento de los propios deseos, creencias, intenciones, o sensaciones, por citar sólo algunos ejemplos de estados mentales. El debate enfrenta a dos intuiciones maestras: por una parte, la intuición según la cual el auto-conocimiento consiste en rastrear, o descubrir, la propia vida mental; por otra parte, la intuición según la cual el auto-conocimiento consiste en la constitución, de manera activa, de la propia vida mental. La primera intuición ha dado lugar a una explicación que podría denominarse *epistémica* del fenómeno del auto-conocimiento, y la segunda a una que podría denominarse *constitutiva*.

El objetivo de este capítulo es proponer una explicación diferente, una alternativa expresivista a las propuestas epistémica y constitutiva que dominan el debate contemporáneo. La intuición maestra detrás de esta alternativa es que el auto-conocimiento es una función de las capacidades expresivas del sujeto en cuestión. A primera vista, parece difícil tejer una explicación alternativa del auto-conocimiento con esos mimbres, pues no es en absoluto obvio que dicha intuición haya de ser negada por quienes defienden una explicación epistémica o una constitutiva. Por eso, la defensa de una concepción expresivista del auto-conocimiento requiere, como primer paso, una formulación precisa de dicha alternativa. Este es el objetivo principal del capítulo: formular, y al tiempo defender mínimamente, una concepción expresivista del auto-conocimiento.

Para ello, se va a proceder del siguiente modo. En primer lugar, se caracterizará de manera precisa la naturaleza del debate contemporáneo sobre el auto-conocimiento, insistiendo tanto en los presupuestos comúnmente aceptados (sección 2), como en las diferencias que enfrentan a los defensores de una explicación epistémica y los de una explicación constitutiva (sección 3). En segundo lugar, se introducirá una distinción entre dos enfoques que se pueden adoptar respecto al auto-conocimiento, uno sincrónico y otro diacrónico (sección 4). A continuación, se vinculará la alternativa expresivista al enfoque diacrónico, frente al enfoque sincrónico reinante en el debate contemporáneo (sección 5). El resultado final será la presentación y defensa de la propuesta expresivista como una alternativa con poder explicativo genuino, que tiene una ventaja respecto a las otras dos explicaciones en liza: a saber, ser una explicación más simple, y por lo tanto preferible *ceteris paribus* (sección 6). El capítulo termina señalando cómo la alternativa expresivista defendida encaja con intuiciones filosóficas relacionadas (sección 7).

2. El problema contemporáneo del auto-conocimiento

El debate contemporáneo sobre el auto-conocimiento concierne la naturaleza de las auto-adcripciones de estados mentales típicos, como deseos, creencias y sensaciones, entre otros. Más en concreto, concierne la naturaleza de las auto-adcripciones de estados mentales típicos, tanto frente a las adcripciones de esos mismos estados mentales a otros, como frente a las auto-adcripciones de estados corporales.¹ Es así porque las auto-adcripciones de estados mentales típicos están sujetas a principios regulativos ausentes de las adcripciones mentales a otros o de las auto-adcripciones corporales. Entre estos principios es común señalar los siguientes: cuando alguien se atribuye a sí mismo (por ejemplo) creencias o deseos particulares, normalmente no se duda de lo que el sujeto dice de sí mismo; ni se le corrige; ni se espera que dé razones a favor de lo que dice. Más aún, si surge una disputa entre terceros acerca de si alguien tiene tal creencia o deseo, se piensa que lo que el sujeto en cuestión dice de sí mismo normalmente resuelve el desacuerdo. En último término, se actúa de este modo porque se concede de entrada que la auto-adcripción del sujeto tiene un estatuto especial: a saber, es verdadera por defecto.

En cambio, en el caso de los demás, la adcripción de estados mentales no está sujeta a los mismos principios regulativos. Así, cuando se atribuyen a otros deseos o creencias particulares, se piensa, al menos en condiciones normales, que es procedente pedir y dar razones a favor de lo que se dice; y normalmente se acepta que lo que se dice pueda ser objeto de duda y de corrección, bien por parte de terceros, bien por los sujetos a los que se les hace la atribución. Algo parecido sucede con las auto-adcripciones de estados corporales, pues en condiciones normales, cuando uno dice de sí mismo (por ejemplo) que mide tanto o pesa cuanto, sus afirmaciones son susceptibles de duda y corrección por otros, si las razones que se dan, cuando se solicitan, no se consideran satisfactorias. Esto es así porque, ni las adcripciones de estados mentales a otros, ni las auto-adcripciones de estados corporales, se consideran verdaderas por defecto.

¹ Típicos, pues puede haber estados mentales propios que no sean fácilmente reconocibles, bien por la complejidad misma de la mente humana, bien por el esfuerzo requerido (por ejemplo, de honestidad con uno mismo). En este capítulo no se va a abordar este aspecto del auto-conocimiento. Pero con ello no se hace trampa, pues el problema contemporáneo del auto-conocimiento concierne los estados mentales típicos.

Que las auto-adcripciones de estados mentales sean verdaderas por defecto no significa que sean infalibles; sólo que, al menos en condiciones normales, se asume que son verdaderas, y por tanto no serán corregidas, a no ser que haya buenas razones adicionales para hacerlo. De este modo, al aceptar el carácter verdadero por defecto de las auto-adcripciones de estados mentales típicos, el debate contemporáneo se distancia desde el inicio de una posición infalibilista acerca del auto-conocimiento (a veces asociada con la concepción cartesiana de la mente). Al mismo tiempo, se distancia también de cierta forma de expresivismo (que a veces se ha atribuido a Wittgenstein; cf. Bar-On y Long, 2001), que rechaza que las auto-adcripciones de estados mentales sean susceptibles de verdad o falsedad, dado su carácter expresivo.

Es importante reparar en la cláusula de normalidad contenida en la descripción del fenómeno: las auto-adcripciones de estados mentales típicos no son objeto de duda o de corrección *en condiciones normales*, ni se espera que el sujeto dé razones a favor de lo que dice acerca de sí mismo *en condiciones normales*, y se les confiere el estatuto de verdaderas por defecto *en condiciones normales*. ¿Pero cuáles son esas condiciones? En la literatura contemporánea se han señalado varias, incluidas la sinceridad (Wright, 2001a, 321; Heal, 1994, 21; Coliva, 2009, 370), el entendimiento (Wright, 2001a, 321; Coliva, 2009, 370), la inteligencia (Shoemaker, 1988, 192), la racionalidad (Shoemaker, 1988, 192; Peacocke, 1998, 84), la posesión de conceptos (Shoemaker, 1988, 192; Peacocke, 1998, 71; Bilgrami, 1998, 209; Coliva, 2009, 370), o la acción responsable (Bilgrami, 1998, 222). La idea central aquí es que el problema filosófico del auto-conocimiento atañe no a la auto-adcripción *simpliciter* de estados mentales típicos, sino a la auto-adcripción en determinadas condiciones, a saber por sujetos que cumplen al menos algunas de las condiciones indicadas en este párrafo.²

Resumiendo, el debate contemporáneo sobre la naturaleza de las auto-adcripciones de estados mentales típicos se produce sobre un trasfondo de acuerdo muy amplio, que incluye no sólo la caracterización del fenómeno (a saber, la existencia de diferencias, en condiciones normales, con las adcripciones de estados mentales a otros y con las auto-adcripciones de estados corporales), sino también el rechazo del infalibilismo y de cierta versión del expresivismo. Sin embargo, como se verá a continuación, este amplio acuerdo es compatible con la existencia de dos propuestas explicativas radicalmente encontradas.

3. Las propuestas epistémica y constitutiva

Según la propuesta epistémica, los principios que, tal como se ha señalado arriba, regulan la auto-adcripción de estados mentales típicos se explican por la existencia de un peculiar mecanismo epistémico, distinto del que sustenta la adcripción de estados mentales a otros y la auto-adcripción de estados corporales. Así, nos atribuimos propiedades corporales a nosotros mismos, como tener un cierto peso o una cierta altura, a partir de la observación de determinados instrumentos de medida. De manera parecida, la atribución de estados mentales a otros descansa, bien en la observación directa de su conducta o sus rasgos corporales, por ejemplo faciales (en el caso de que la vida mental ajena sea objeto directo de percepción en la conducta o en los rasgos corporales), bien en la observación de su conducta y sus rasgos corporales más la inferencia al estado mental asociado a los primeros (en el caso de que la

² Llegados a este punto, pues, y para proseguir con nuestra investigación, no es necesario dirimir si todas y cada una de ellas son condiciones necesarias para el auto-conocimiento, siempre y cuando se convenga con el planteamiento contemporáneo del debate que algunas lo son. Por ello, en el resto del capítulo habrá de darse por sentado que la cláusula de normalidad está activada.

vida mental ajena no sea objeto directo de percepción). En cualquier caso, tanto la auto-adscripción de estados corporales como la adscripción de estados mentales a otros descansan sobre el ejercicio de capacidades perceptivas. Sin embargo, parece obvio que la auto-adscripción de estados mentales típicos ni requiere, ni se basa normalmente, en la observación, o en general la percepción, de uno mismo, en concreto de la propia conducta o de los propios rasgos corporales. Más bien, decimos de nosotros mismos que sentimos tal o cual emoción, o que albergamos tal o cual deseo o creencia, antes, e incluso con independencia, de la existencia de manifestaciones corporales perceptibles.

Pero, ¿cuál es exactamente el mecanismo epistémico que subyace a la auto-adscripción de estados mentales típicos? Aunque la idea que nos puede venir a mentes de manera inmediata sea algo así como la “observación” (o cuasi-observación) no sensorial del estado mental en cuestión, entendido como algo distinto de la manifestación corporal perceptible sensorialmente, ésta no es la opción elegida por los defensores contemporáneos de la propuesta epistémica. La razón es que en dicha idea resuenan los ecos del dualismo ontológico mente-cuerpo, una tesis generalmente rechazada en la filosofía contemporánea. De hecho, no hay una única propuesta epistémica, que pudiera considerarse estándar en el debate contemporáneo, sino una variedad de propuestas para caracterizar el mecanismo epistémico subyacente a las auto-adcripciones de estados mentales típicos. He aquí algunos ejemplos:

Si asentimos a la identificación materialista de los estados mentales con estados materiales del cerebro, podemos decir que la introspección es un proceso de auto-escaneado del cerebro. (Armstrong, 1968, 324)

Llamemos “auto-adscripción fundamentada conscientemente” a la adscripción que un sujeto se hace a sí mismo de una actitud con un contenido determinado, por la razón de que tiene una actitud consciente concomitante con ese mismo contenido. En el ejemplo propuesto, la auto-adscripción de la creencia de que Dubček era primer ministro durante la invasión de la Unión Soviética [de Checoslovaquia] es una auto-adscripción fundamentada conscientemente [a saber, en el juicio de que Dubček era el primer ministro entonces]. [...] Auto-adscribirse una creencia porque está basada en un estado consciente es, por supuesto, una de las maneras, propias de la auto-adscripción, en las que un sujeto puede adquirir conocimiento auto-adscribiéndose una creencia. (Peacocke, 1998, 72) En esencia, mi propuesta es que, cuando me formo la creencia de que tengo una creencia determinada, normalmente me formo la creencia de orden superior a partir del estado que constituye la base de mi creencia de primer orden. Lo que hago es, por así decirlo, mirar más allá de la creencia con el fin de auto-adscribirmela. Por ejemplo, cuando creo que, en este mismo momento, creo que hay una manzana enfrente de mí, me formo esa creencia a partir de la base que tengo para creer que estoy enfrente de una manzana (una base tal sería una experiencia perceptiva de una manzana). [...] La idea central es, pues, que nos auto-adscribimos creencias a partir de la base que tenemos a favor de esas creencias. (Fernández, 2013, 49)³

Lo que es común a todas estas propuestas (es decir, lo que las convierte en propuestas epistémicas) es, por una parte, que hay un mecanismo epistémico para rastrear, o descubrir, la propia vida mental, y por otra, que el estado de cosas rastreado precede, o al menos es independiente, de la auto-adscripción. Ciertamente, hay diferencias en la concepción del mecanismo epistémico de rastreo, pues algunos lo conciben como un sistema de escaneado interno del propio cerebro (Armstrong), mientras que para otros es simplemente la atención

³ Fernández dice algo parecido para explicar la auto-adscripción de deseos. Cf. 2013, 79-108.

consciente a los propios pensamientos y actitudes mentales cuando se producen (Peacocke), y aun para otros consiste en la búsqueda en el mundo de la base de las propias actitudes, como deseos y creencias (Fernández). Pero, de una forma u otra, todos aceptan que el auto-conocimiento es una forma de acceso a un estado de cosas psicológico independiente de la auto-adscripción.

Más aún, todos apelan a ese mecanismo de acceso a la propia mente para explicar las peculiaridades de la auto-adscripción de estados mentales señaladas antes. Así, uno no puede proceder al escaneado *interno* del cerebro ajeno, ni es la auto-adscripción de propiedades corporales como el peso o la altura el resultado de tal proceso de escaneado interno. Análogamente, uno no puede atender conscientemente a los pensamientos y las actitudes ajenas como lo hace en su propio caso; ni puede hacerlo con los propios estados corporales (más bien, como hemos visto, recurre a sus capacidades perceptivas). Finalmente, el contenido de los estados mentales ajenos se fija mirando, no al mundo tal como se le presenta a uno en sus experiencias perceptivas (por ejemplo), sino a la conducta y los rasgos corporales de los otros sujetos. Es, por lo tanto, la existencia de un distintivo mecanismo epistémico lo que explicaría la naturaleza de las auto-adcripciones de estados mentales típicos, con sus características propias.

La propuesta epistémica, en sus distintas variedades, es ciertamente una explicación muy popular de la naturaleza de las auto-adcripciones de estados mentales típicos. Pero no es la única: igualmente popular es la propuesta constitutiva. Según ésta, la razón por la que la auto-adscripción de estados mentales típicos se regula según los principios señalados arriba, y se distingue por tanto de la adscripción de estados mentales a otros y de la auto-adscripción de estados corporales, no radica en la existencia de un peculiar mecanismo de acceso a la propia mente. Radica, más bien, en el hecho de que en condiciones normales la auto-adscripción trae consigo, o produce, el estado mental en cuestión. Así, en esas condiciones, cuando alguien se pregunta por su propia vida mental, y dice de sí mismo (bien pública, bien silenciosamente) que, por ejemplo, alberga tal o cual deseo o intención, o que siente tal o cual emoción o sensación, su auto-adscripción produce el estado de cosas psicológico que hace verdadera la auto-adscripción. He aquí algunas citas ilustrativas:

La autoridad conferida en condiciones normales a las creencias o a las adcripciones expresadas por un sujeto respecto de sus propios estados intencionales es la de un *principio constitutivo*: no es una consecuencia de la naturaleza de esos estados más una relación epistémica privilegiada del sujeto con esos estados, sino que es un elemento primitivo de las condiciones de identificación de lo que el sujeto cree, espera o tiene la intención de hacer. (Wright, 1989, 632; cursivas en el original)⁴

El lazo constitutivo que vemos en nuestra práctica [de auto-adscripción de creencias] consiste en que se vincula pensar que creo que p (es decir, aquello en lo que consiste la sinceridad de mi afirmación) y creer que p. [...] Resumiendo, estoy justificado cuando me pronuncio sobre mis creencias, no porque tenga un acceso epistémico privilegiado a un estado independiente, sino porque cuando pienso que creo que p, en virtud de ese pensamiento creo que p. (Heal, 1994, 21-22)

⁴ Aunque en esta cita Wright restringe la propuesta constitutiva a las adcripciones (o creencias) sobre los propios estados intencionales, posteriormente la ha generalizado a todas las adcripciones de estados mentales propios. Cf. Wright, 2001b, 368.

Podría glosarse la propuesta constitutiva diciendo que el sujeto es el autor, no el cronista, de su propia vida mental. Así, la auto-adscripción de estados mentales típicos no rastrea un estado de cosas psicológico independiente, sino que lo conforma. Y es por ello por lo que las auto-adcripciones de estados mentales típicos se regulan por los principios mencionados arriba. En la medida en que uno no conforma los estados mentales ajenos, o los estados corporales propios, al adscribirlos, las adcripciones de estados mentales a otros y las auto-adcripciones de estados corporales no cuentan como verdaderas por defecto, y están sometidas a la práctica de petición y dación de razones, y sujetas a la duda y la corrección.

En resumen, pues, el debate contemporáneo sobre el auto-conocimiento está dominado por dos propuestas encontradas: una basada en la existencia de un mecanismo diferenciado para acceder a los estados de cosas psicológicos objetos de auto-adscripción, y otra basada en la constitución de dichos estados de cosas mediante la auto-adscripción. Ambas parecen dar cuenta satisfactoria del fenómeno a explicar: a saber, las diferencias entre los principios que regulan nuestra práctica de auto-adscripción de estados mentales típicos, y los que regulan la adscripción de estados mentales a otros y la auto-adscripción de estados corporales. En ese caso, si ninguna sobresale en razón de su mayor poder explicativo, ¿cómo se ha de proceder para decidir cuál de las dos es la mejor explicación de nuestro *explanandum*? Antes de responder, conviene evaluar un supuesto implícito en la pregunta: a saber, que es una de estas dos propuestas la que ha de proporcionar la mejor explicación del problema filosófico contemporáneo del auto-conocimiento.

4. Los enfoques sincrónico y diacrónico

Aunque, como se ha señalado, las propuestas epistémica y constitutiva proporcionan explicaciones encontradas de las peculiaridades de las auto-adcripciones de estados mentales típicos, es posible encontrar parecidos significativos entre ambas. Así, dado que uno de los datos a explicar es que las auto-adcripciones mentales tienen el estatuto de verdaderas por defecto, los defensores de una y otra propuesta (o sus variedades) se afanan por establecer un vínculo entre las auto-adcripciones mismas y los estados de cosas psicológicos que las hacen verdaderas. O por mejor decir, se afanan por clarificar la naturaleza de ese vínculo, entendido bien en términos de rastreo (o descubrimiento), bien en términos de constitución (o producción, o determinación). Es decir, o las auto-adcripciones de estados mentales típicos son verdaderas (por defecto) porque rastrean el estado de cosas psicológico que las hacen verdaderas, o lo son porque conforman (determinan) el estado de cosas psicológico que las hacen verdaderas.

Al mismo tiempo, dado que las auto-adcripciones mentales que tienen interés filosófico son las realizadas por sujetos que satisfacen determinadas condiciones de sinceridad, inteligencia, racionalidad o sofisticación conceptual (por recordar algunas de las citadas en la sección 2), se tiende a asumir, más o menos explícitamente, que el objeto de estudio son las auto-adcripciones realizadas por los sujetos que satisfacen dichas condiciones. De hecho, las condiciones son satisfechas por los sujetos humanos adultos. En consecuencia, los defensores de una y otra propuesta (o sus variedades) se afanan por clarificar la naturaleza del vínculo entre las auto-adcripciones realizadas por sujetos humanos adultos y los estados de cosas psicológicos que las hacen verdaderas.

Podría denominarse *sincrónico* a este enfoque del problema filosófico del auto-conocimiento. Sincrónico, porque el problema se plantea solamente una vez asumida la existencia de sujetos que satisfacen determinadas condiciones de sinceridad, racionalidad, etc., y se auto-adscriben estados mentales. Es decir, se asume de entrada que los sujetos cuyas auto-adcripciones son objeto de estudio satisfacen determinadas condiciones, y a partir de entonces (y sólo entonces) parece plantearse un problema filosóficamente interesante acerca de cuál es la naturaleza del vínculo entre las auto-adcripciones de dichos sujetos *en un momento dado* y los estados de cosas psicológicos que las hacen verdaderas.⁵

Por contraposición, en un enfoque diacrónico el foco de atención no es la pregunta por la naturaleza del vínculo, en un momento dado, entre las auto-adcripciones mentales de los sujetos que satisfacen las condiciones mencionadas y los estados de cosas psicológicos que las hacen verdaderas, sino la pregunta por cómo se ha llegado ahí. Dicho de otro modo, en un enfoque diacrónico se presta especial atención a la adquisición de la capacidad de auto-adcripción de estados mentales típicos por parte de los sujetos humanos adultos, es decir los sujetos que satisfacen las condiciones mencionadas.

Nótese que los enfoques sincrónico y diacrónico son justamente eso, enfoques distintos del mismo problema, pues con el cambio de enfoque el problema no cambia. El problema sigue siendo la explicación de las peculiaridades de las auto-adcripciones de estados mentales típicos, frente a la adcripción de estados mentales a otros y a la auto-adcripción de estados corporales. Lo que cambia es que, mientras en el enfoque sincrónico el éxito explicativo recae sobre la clarificación del vínculo entre la auto-adcripción mental por parte de los sujetos humanos adultos en un momento dado y el estado de cosas psicológico que la hace verdadera, en el enfoque diacrónico el éxito explicativo recae sobre la existencia de una historia plausible que contar acerca de la adquisición de la capacidad de auto-adcripción de estados mentales típicos por dichos sujetos.

Nótese también que las propuestas epistémica y constitutiva, aunque adoptan un enfoque sincrónico, no niegan la existencia de una historia plausible de la adquisición de la capacidad de auto-adcripción de estados mentales típicos. Más bien, lo que hacen es asumir sin más la existencia de dicha historia, mientras ponen el foco explicativo en la naturaleza, bien epistémica, bien constitutiva, del vínculo entre las auto-adcripciones particulares de estados mentales típicos en un momento dado y los estados de cosas psicológicos que las hacen verdaderas.

Asumiendo, pues, que estos son dos enfoques posibles, la pregunta pertinente es: ¿podría un enfoque diacrónico ofrecernos nuevas pistas para encarar el problema filosófico del auto-conocimiento? O mejor, ¿podría ofrecer una alternativa explicativa a las propuestas epistémica y constitutiva que dominan el debate contemporáneo?

⁵ Esta formulación del problema vuelve especialmente problemática la naturaleza del auto-conocimiento basado en la memoria; es decir, el vínculo entre la auto-adcripción de un estado mental recordado y el estado de cosas psicológico que la hace verdadera. ¿Se trata de un caso de producción (en el presente) de un estado de cosas psicológico pasado? ¿Se trata de un caso de rastreo de un estado de cosas psicológico pasado, es decir inexistente en el momento de la auto-adcripción, como si algo que no es el caso pudiera hacer verdadera la auto-adcripción? ¿O se trata acaso del rastreo, no de un estado de cosas psicológico pasado, sino de uno existente en el presente (el recuerdo como tal), como si la memoria nos diera acceso al presente?

5. Una alternativa expresivista

La historia que es plausible contar acerca de la adquisición de la capacidad de auto-adscripción de estados mentales típicos es una historia ontogenética; es decir, una historia acerca del desarrollo psicológico individual del niño hasta alcanzar la edad adulta. Durante ese período, y de manera progresiva, el niño adquiere una mente adulta, así como la capacidad para auto-adscribirse sus propios estados mentales. Dicho de otro modo, el niño adquiere la capacidad para expresar su propia vida mental. ¿Cómo ha de entenderse esto en el marco del problema contemporáneo sobre el auto-conocimiento, de tal manera que no sea algo banal e insulso?

Por una parte, podría entenderse como la idea de que el niño adquiere las herramientas lingüísticas para hacer pública su vida mental, como si ésta estuviera ya configurada antes de la adquisición del lenguaje, y lo que la adquisición del lenguaje le permite es dársela a conocer a los demás. Aquí, la historia ontogenética se entiende como la adquisición de la capacidad de expresar, en el sentido de dar a conocer a los demás, la propia vida mental, mediante el recurso lingüístico de la auto-adscripción. Ciertamente, hay un poso de verdad en esta manera de entender el desarrollo psicológico individual, puesto que es innegable que damos a conocer nuestra propia vida mental a los demás al auto-adscribérnosla. Sin embargo, la historia ontogenética que es plausible contar no puede reducirse a eso, so pena de ignorar el hecho de que durante el desarrollo psicológico individual se configura, de manera progresiva, la mente adulta misma, que no está dada de antemano.

Por ello, parece más acertado entender el desarrollo psicológico individual como el proceso de adquisición progresiva de una mente adulta, gracias a las posibilidades de interacción con la comunidad que se le abren al niño al adquirir el lenguaje. Pero, a su vez, esto puede entenderse de dos maneras distintas. Según una de ellas, el desarrollo psicológico individual se entiende como el doble proceso mediante el cual (i) se adquiere de manera progresiva una mente adulta, y simultáneamente (ii) se adquiere la capacidad de expresar (léase dar a conocer a los demás) la propia vida mental mediante la auto-adscripción. Según esto, en ambos procesos juega un papel central la adquisición del lenguaje, pero lo importante es que se trata de procesos distintos: una cosa es tener los contenidos mentales (por ejemplo, deseos, creencias o sensaciones) propios de la edad adulta, y otra cosa es adscribírselos a uno mismo. Nuevamente, hay aquí un poso de verdad, pues un sujeto adulto normal puede dejar de auto-adscribirse los estados mentales que de hecho tiene, y de ese modo no hacerlos públicos (al menos mediante el lenguaje). Sin embargo, la historia ontogenética que es plausible contar no puede terminar así, so pena de convertir en un hecho contingente bruto que los sujetos adultos normales (léase, los que han adquirido una mente adulta durante el proceso de desarrollo ontogenético) poseen la capacidad de adscribirse sus propios estados mentales. Así lo muestra el siguiente test: si un sujeto al que se le presupone una mente adulta normal resultara carecer de la capacidad de auto-adscribirse estados mentales típicos, se revisaría de inmediato dicha presuposición.

La alternativa es entender el desarrollo psicológico individual como un proceso consistente en la adquisición progresiva de una mente adulta mediante la adquisición de capacidades expresivas, incluida la de auto-adscripción de estados mentales. Según esto, conforme el niño participa en interacciones cada vez más complejas con su comunidad gracias al lenguaje, incluida la adscripción a sí mismo de distintos estados mentales, se va configurando progresivamente su mente adulta. Conviene entender bien lo que se está diciendo aquí. En primer lugar, no se está diciendo que el niño carece de vida mental al comienzo, y que ésta

surge, por así decirlo de la nada, durante el proceso de desarrollo individual descrito. Más bien, el recién nacido ya tiene estados mentales desde el inicio, por más que sean limitados y simples, y durante el proceso de desarrollo individual se transforman en los estados mentales de la edad adulta.

En segundo lugar, se está diciendo que el proceso de desarrollo individual está tutelado, por así decirlo, por los miembros adultos normales de la comunidad, que supervisan las interacciones (también las lingüísticas) con el niño. Por lo tanto, la formación de la mente adulta en el niño es un proceso fundamentalmente social. A modo de contraste, podría decirse que los individuos que no participan en dicho proceso, bien por razón de su aislamiento, bien por razón de sus limitaciones cognitivas (congénitas o no), carecen de una mente adulta plena.

En tercer lugar, se está corrigiendo la idea según la cual la capacidad de auto-adscripción de estados mentales es posterior, o quizás simultánea, pero en todo caso distinta, a la posesión de una mente adulta. No es posterior, pues el niño adquiere una mente adulta, gracias a la interacción lingüística con su comunidad, incluida la capacidad de decir de sí mismo (por ejemplo) que siente tal o cual sensación o emoción, o alberga tal o cual deseo o creencia. Pero tampoco es meramente simultánea, pues no es un hecho contingente bruto que los sujetos con una mente adulta normal poseen la capacidad de auto-adscribirse sus estados mentales, como se ha señalado antes. Por lo tanto, ni la posesión de una mente adulta precede a la capacidad de auto-adscripción de estados mentales, ni es la primera simultánea pero independiente de la segunda. Al contrario, según la alternativa expresivista ahora expuesta, existe una relación muy estrecha entre la adquisición de una mente adulta y la adquisición de la capacidad de auto-adscripción de estados mentales típicos. El lazo de unión entre ambas es el proceso de desarrollo individual, durante el cual se adquiere una mente adulta gracias, al menos en parte, a la adquisición de la capacidad de atribuirse a sí mismo estados mentales típicos.⁶

Puede resultar difícil aceptar esta concepción de la adquisición de la mente adulta. Sin negar la existencia de un proceso de desarrollo psicológico individual, puede parecer que las claves de dicho proceso están no, o no primordialmente, en la rica relación lingüística con los miembros adultos normales de la comunidad, sino en el desarrollo de una compleja estructura neuronal, un proceso biológico que se rige por leyes o mecanismos muy distintos a los descritos en los párrafos precedentes. Es sin duda obvia la existencia de una base neuronal de la mente humana, y la historia ontogenética contada arriba de ningún modo la niega. Así, la adquisición de una mente adulta requiere que determinadas estructuras neuronales (así como otras estructuras corporales) funcionen adecuadamente (véase lo dicho antes sobre los individuos con limitaciones cognitivas, congénitas o no). Pero dicho eso, la historia ontogenética contada insiste en que la mente adulta se adquiere durante un complejo proceso de relación con la comunidad, gracias al lenguaje. El sesgo neuronal (la idea de que la mente es idéntica, o reducible, a una estructura neuronal) puede ciertamente ser un obstáculo para aceptar esa historia ontogenética, pero en ese caso tal vez habría que indagar en la idoneidad de esa concepción de la mente (algo que no puede hacerse aquí).

⁶ Al menos en parte, porque el proceso de adquisición de una mente adulta en el seno de una comunidad, bajo la tutela de miembros adultos normales de la misma, es muy rico y altamente complejo, e incluye también la adscripción y el reconocimiento de estados mentales típicos en otros.

Sea como fuere, y avanzando en nuestra investigación, la pregunta clave es cuál es la contribución de la historia ontogenética contada al problema contemporáneo del auto-conocimiento. Según se ha explicado, se trata del problema de explicar las diferencias entre las auto-adcripciones de estados mentales típicos en condiciones normales (de sinceridad, inteligencia, racionalidad, etc.), por una parte, y las adcripciones de estados mentales a otros y las auto-adcripciones de estados corporales, por otra. En concreto, lo que hay que explicar es por qué sólo las primeras (pero no las segundas) se regulan por principios tales como no estar normalmente sujetas a duda y corrección, por ser ajenas a la práctica de pedir y dar razones, por servir normalmente para resolver disputas entre terceros, y por conferírseles el estatuto de verdaderas por defecto. Pues bien, una historia ontogenética como la contada permite dar una explicación del asunto. La clave es que un aspecto central del proceso de desarrollo ontogenético individual es que la adquisición de una mente adulta y la adquisición de la capacidad de auto-adcripción de estados mentales típicos van de la mano (aunque no en el sentido de dos procesos meramente simultáneos, como se ha explicado arriba). Esto quiere decir que el proceso de desarrollo ontogenético en el que toma forma la mente adulta es al mismo tiempo un proceso de adquisición de habilidades expresivas (o auto-expresivas, si se quiere). Y sobre esta idea se puede construir una concepción expresivista del auto-conocimiento.

Para empezar, si el proceso de desarrollo psicológico individual es también un proceso de adquisición de habilidades expresivas, se explica por qué hay sujetos que se auto-adscriben estados mentales típicos en determinadas condiciones de sinceridad, inteligencia, racionalidad, etc. Hasta ahora, en el planteamiento del problema del auto-conocimiento (sección 2), y en la discusión de las propuestas epistémica y constitutiva (sección 3), se tomaba eso como el hecho bruto que había que explicar. Pero ahora resulta no ser un simple hecho bruto, pues descansa sobre un hecho ontogenético ulterior: a saber, la existencia de un proceso en el que los individuos adquieren la madurez de la edad adulta (incluidas la sinceridad, la inteligencia, la racionalidad, etc.), al tiempo que adquieren las capacidades expresivas de auto-adcripción de estados mentales típicos. Ciertamente, esto no es una explicación como tal del fenómeno del auto-conocimiento, pero tampoco es algo vacuo, pues sirve para arrojar luz adicional sobre el fenómeno mismo.

Además, si el proceso de desarrollo psicológico individual es también un proceso de adquisición de habilidades expresivas, se explican las diferencias señaladas entre las auto-adcripciones de estados mentales típicos, por una parte, y las adcripciones mentales a otros y las auto-adcripciones corporales, por otra. Dado que la formación de la mente adulta va de la mano de la adquisición de habilidades expresivas, una vez que se han adquirido dichas habilidades expresivas, se pueden seguir ejercitando durante la edad adulta. Así, en condiciones normales, un sujeto adulto puede auto-adscribirse sus estados mentales típicos, justamente porque posee las habilidades expresivas (auto-expresivas) que ha adquirido durante el proceso de desarrollo individual. En esas condiciones, lo que el sujeto dice de sí mismo tiene el estatuto de verdadero por defecto, y no se somete a duda o corrección, ni se le exige al sujeto que dé razones a favor de lo que dice, simplemente porque se le toma por un sujeto adulto normal, que ha pasado satisfactoriamente por el proceso de desarrollo individual, y por tanto ha adquirido las habilidades expresivas que sigue ejercitando después.

Muy distintas son las cosas en el caso de la auto-adcripción de estados corporales, y de la adcripción de estados mentales a otros. Por una parte, estados corporales como tener un cierto peso o una cierta altura no tienen asociados a ellos capacidades expresivas (auto-expresivas) semejantes; por lo tanto, no hay nada en el proceso de desarrollo individual que

implique que lo que uno dice, por ejemplo, acerca de su peso o su altura tenga un estatuto epistémico especial, o sea verdadero por defecto. Eso es así incluso si uno es un sujeto adulto normal que ha pasado satisfactoriamente por el proceso de desarrollo individual: al fin y al cabo, dicho proceso no implica la adquisición de habilidades para expresar, sin más (es decir, sin recurrir a instrumentos de medida), cuál es el propio peso o la propia altura.

Por otra parte, aunque la atribución de estados mentales a otros es un aspecto central del proceso de desarrollo individual (dado el carácter social del mismo, tal como se ha señalado antes), una cosa son las habilidades expresivas (auto-expresivas) adquiridas durante dicho proceso, y otra son las habilidades para reconocer estados mentales en otros. Más aún, las habilidades de reconocimiento y de atribución de estados mentales ajenos se adquieren en contextos donde la duda, la corrección, la petición y dación de razones son prominentes, y en los que no lo es tomar tal atribución como verdadera por defecto. Es esta diferencia de habilidades, pues, la que explica que sólo se confiera un estatuto especial al ejercicio de las habilidades expresivas (auto-expresivas), en la auto-adscrición de estados mentales.

En resumen, pues, es posible defender la existencia de una concepción expresivista del auto-conocimiento, según la cual éste consiste en el ejercicio de las habilidades expresivas adquiridas durante el proceso de desarrollo individual en el que se forma la mente adulta. Más aún, esta concepción expresivista es una alternativa a las propuestas epistémica y constitutiva en el debate contemporáneo, ya que no se apela ni a mecanismos para rastrear estados de cosas psicológicos independientes, ni a la constitución de la propia vida mental mediante la misma auto-adscrición. Podrían glosarse estas diferencias diciendo que, según la alternativa expresivista, la autoridad de la primera persona no es la autoridad del cronista, ni del autor, de la propia vida mental, sino la autoridad de la madurez.

Ahora bien, si se acepta que existen estas tres opciones en liza, y que cada una de ellas permite dar cuenta del fenómeno que se desea explicar, ¿qué explicación es mejor?

6. Simplicidad explicativa

Cuando existen varias explicaciones posibles, pero incompatibles entre sí, de uno y el mismo fenómeno, es legítimo optar por la explicación que requiere menos recursos. La simplicidad es, pues, un valor explicativo: no uno absoluto, pero sí uno al que es legítimo recurrir si otros valores han sido satisfechos (por ejemplo, si se salva el *explanandum*). Esta consideración nos permite desplegar el siguiente argumento pro-expresivista: la alternativa expresivista es la mejor explicación del auto-conocimiento (de las tres consideradas), pues aunque todas explican el *explanandum* (el tratamiento distintivo de las auto-adscriciones de estados mentales típicos frente a las adscriciones de estados mentales a otros y a las auto-adscriciones de estados corporales), la alternativa expresivista ofrece la explicación más simple.

En las páginas anteriores se ha mostrado que las propuestas epistémica, constitutiva y expresivista consiguen, cada una a su manera, explicar las características distintivas de las auto-adscriciones de estados mentales típicos. Por lo tanto, para defender el argumento pro-expresivista anterior sólo resta mostrar ahora que la alternativa expresivista es la más simple de las tres. La consideración crucial en este sentido estriba en algo que ya se ha señalado de pasada antes (sección 4): a saber, que las propuestas epistémica y expresivista no niegan que haya una historia que contar acerca de la adquisición de la capacidad de auto-adscrición de estados mentales típicos. Más bien, lo que sucede es que ponen el foco explicativo no ahí, sino en la naturaleza, bien epistémica, bien constitutiva, del vínculo entre las auto-

adscripciones de estados mentales típicos en un momento dado y los estados de cosas psicológicos que las hacen verdaderas. Así pues, las explicaciones epistémica y constitutiva del auto-conocimiento explotan dos tipos de recursos: los que tienen que ver con la existencia y adquisición de habilidades expresivas (auto-expresivas), y los que tienen que ver con los mecanismos particulares (bien epistémicos, bien constitutivos) mediante los cuales el ejercicio de dichas habilidades expresivas (en determinadas condiciones) en un momento dado resulta en auto-adscripciones que son verdaderas por defecto. En cambio, la alternativa expresivista explota sólo los recursos que tienen que ver con la existencia de las habilidades expresivas (auto-expresivas) adquiridas durante el proceso de desarrollo psicológico individual. Por consiguiente, la alternativa expresivista es la más simple de las tres en liza a la hora de explicar el problema contemporáneo del auto-conocimiento.

Nótese que, según la historia ontogenética contada en estas páginas, la adquisición de las habilidades expresivas (auto-expresivas) durante el proceso de desarrollo psicológico individual no es el fruto de mecanismos de rastreo o de constitución de la propia vida mental en momentos dados sucesivos. No se trata, pues, de que el proceso de desarrollo psicológico consista en aprender a rastrear o constituir la propia vida mental. Si así fueran las cosas, la concepción expresivista no sería una auténtica alternativa a las propuestas epistémica y constitutiva, pues dependería de la existencia de mecanismos epistémicos o constitutivos. Pero las cosas no son así, tal como se ha explicado arriba. Por una parte, el desarrollo psicológico individual no consiste en el acceso, en un momento dado, a un estado de cosas psicológico independiente, sino en la formación misma de la vida mental adulta. Por otra parte, no hay ya una capacidad de auto-adscripción cuyo ejercicio en un momento dado conforma la propia vida mental, sino que justamente se está adquiriendo dicha capacidad. Por eso, no puede objetarse al argumento pro-expresivista anterior, basado en la simplicidad explicativa, que la alternativa expresivista no es la más simple, pues recurre (quizás veladamente) a mecanismos epistémicos o constitutivos. No lo hace, tal como se expuso en la sección anterior, y se acaba de recordar.

Por supuesto, el argumento basado en la simplicidad explicativa no es un argumento definitivo a favor de la concepción expresivista del auto-conocimiento. Como se ha señalado, la simplicidad no es un valor absoluto, sino uno dependiente de la existencia de otros valores. De momento, la alternativa expresivista emerge victoriosa porque tiene el valor adicional de la simplicidad: es decir, porque da cuenta del fenómeno que ha de ser explicado, y lo hace de la manera más simple. No obstante, una defensa más enérgica de la concepción expresivista debería mostrar que, además de dar cuenta del *explanandum* y hacerlo de manera más simple que otras propuestas, es coherente con intuiciones filosóficas relacionadas que merecen ser respetadas. Así, la simplicidad no puede ser una coartada para renunciar a la coherencia general de la explicación propuesta. A continuación, y antes de terminar, se va a mostrar que la concepción expresivista no defrauda en este sentido.

7. Expresión, verdad y sinceridad

Se ha señalado en la sección 2 que el planteamiento contemporáneo del problema del auto-conocimiento se distancia de cierta forma de expresivismo que rechaza que las auto-adscripciones de estados mentales típicos sean susceptibles de verdad o falsedad. En consonancia con esto, en los últimos años algunos autores han defendido que las auto-adscripciones son expresiones y, a la vez, son susceptibles de verdad y falsedad (Finkelstein, 2003; Bar-On, 2004). La idea que subyace a esta defensa es que las auto-adscripciones de

estados mentales típicos son expresiones lingüísticas, y *qua* expresiones *lingüísticas* pueden ser verdaderas o falsas. Por supuesto, al mismo tiempo ha de reconocerse que *qua* expresiones, las auto-adcripciones de estados mentales típicos son susceptibles de sinceridad o insinceridad. El resultado es que, en esta defensa del expresivismo, las auto-adcripciones de estados mentales típicos son, a la vez, susceptibles de sinceridad o insinceridad (*qua* expresiones), y susceptibles de verdad y falsedad (*qua* expresiones lingüísticas).

Este resultado parece intuitivo, pero los casos de desacuerdo plantean el siguiente problema. Si yo me auto-adscribo un determinado estado mental, y tú me contradices, ¿es esto un desacuerdo acerca de la verdad de lo que digo de mí mismo, o es un desacuerdo acerca de la sinceridad de mi expresión? Parece claro que, como se trata de una expresión lingüística, el desacuerdo tiene que ver con la verdad de lo que digo de mí mismo: al contradecirme, estás negando que lo que yo decía sea verdadero. Pero, ¿es o no también un desacuerdo acerca de la sinceridad de mi expresión? Si lo es, el expresivismo contemporáneo nos debe una explicación de por qué es así, especialmente dado que distingue entre las condiciones de sinceridad de una expresión y sus condiciones de verdad (*qua* expresión lingüística). En cambio, si no lo es, el expresivismo contemporáneo nos debe un modelo que nos permita reconocer la existencia de casos de desacuerdo acerca de la sinceridad de una auto-adcripción (lingüística) de un estado mental típico.⁷

En esta disyuntiva, la alternativa expresivista defendida arriba nos ofrece una salida: a saber, en el caso de los sujetos humanos adultos normales, que tienen las capacidades lingüísticas para auto-adcribirse la propia vida mental, verdad y sinceridad no son propiedades distintas de las auto-adcripciones de estados mentales típicos. Así, lo que sucede durante el proceso de desarrollo psicológico individual es que se forma la mente adulta, al tiempo que se adquiere la capacidad para decir sinceramente de uno mismo que (por ejemplo) se tiene tal emoción o sensación, o que se albergan unos deseos y creencias u otros. Es decir, la adquisición de la capacidad de expresión (auto-expresión) sincera y la fijación de la verdad del caso (el estado de cosas psicológico que hace verdadera la auto-adcripción) están íntimamente ligadas por la naturaleza del desarrollo psicológico individual. En consecuencia, en los casos de desacuerdo mencionados, cuestionar la verdad (contradecir la auto-adcripción primera) es también cuestionar la sinceridad del sujeto, dadas las condiciones del desarrollo psicológico individual. Tenemos, pues, el modelo expresivista demandado en el párrafo anterior, para los casos de desacuerdo respecto a auto-adcripciones de estados mentales típicos.

En resumen, pues, la alternativa expresivista presentada y defendida en este capítulo no sólo es, de las tres propuestas consideradas, la explicación más simple del fenómeno del auto-conocimiento (el tratamiento distintivo de las auto-adcripciones de estados mentales típicos), sino que además permite iluminar la relación entre las nociones de expresión, verdad y sinceridad, que plantea problemas a otras defensas contemporáneas del expresivismo en el auto-conocimiento. Es, por lo tanto, una opción que se ha hecho merecedora de una mayor atención en el debate contemporáneo.

⁷ En García Rodríguez (2012) se exploran los problemas que los casos de desacuerdo como éste provocan al expresivismo de Finkelstein y Bar-On.

Bibliografía

- Armstrong, David M. (1968), *A materialist theory of the mind*, London, Routledge.
- Bar-On, Dorit (2004), *Speaking my mind*, Oxford University Press, Oxford.
- Bar-On, Dorit y Long, Douglas C. (2001), “Avowals and first-person privilege”, *Philosophy and phenomenological research*, vol. LXII, pp. 311-335.
- Bilgrami, Akeel (1998), “Self-knowledge and resentment”, en C. Wright, B. Smith y C. Macdonald (eds.) *Knowing our own minds*, Clarendon Press, Oxford, pp. 207-41.
- Coliva, Annalisa (2009), “Self-knowledge and commitments”, *Synthese*, vol. 171/3, pp. 365-75.
- Fernández, Jordi (2013), *Transparent minds*, Oxford University Press, Oxford.
- Finkelstein, David (2003), *Expression and the inner*, Harvard University Press, Cambridge, Mass. y London.
- García Rodríguez, Ángel (2012), “How to be an expressivist about avowals today”, *Nordic Wittgenstein Review*, vol. 1, pp. 81-101.
- Heal, Jane (1994), “Moore’s paradox: a Wittgensteinian approach”, *Mind*, vol. 103, pp. 5-24.
- Peacocke, Christopher (1998), “Conscious attitudes, attention and self-knowledge”, en C. -Wright, B. Smith y C. Macdonald (eds.) *Knowing our own minds*, Clarendon Press, Oxford, pp. 63-98.
- Shoemaker, Sydney (1988), “On knowing one’s own mind”, *Philosophical perspectives*, vol. 2, pp. 183-209.
- Wright, Crispin (1989), “Wittgenstein’s later philosophy of mind: sensation, privacy and intention”, *The journal of philosophy*, vol. 86, pp. 622-34.
- Wright, Crispin (2001a), “The problem of self-knowledge (I)”, en *Rails to infinity*, Harvard University Press, Cambridge, Mass. y London, pp. 319-44.
- Wright, Crispin (2001b), “The problem of self-knowledge (II)”, en *Rails to infinity*, Harvard University Press, Cambridge, Mass. y London, pp. 345-73.