

LA CONCIENCIA ERRÓNEA. DE SÓCRATES A TOMÁS DE AQUINO*

ALEJANDRO G. VIGO**

Resumen: En el ámbito de la acción moral, el principio socrático de que *nadie yerra voluntariamente* implica que toda vez que un agente elige algo lo hace por considerarlo, al mismo tiempo, como bueno o, al menos, preferible a otra cosa: su elección es internamente racional. La tesis socrática sobre la conexión estructural entre error y autoengaño constituye, sin duda, uno de los aportes más decisivos al pensamiento filosófico occidental. De esta concepción en torno a la naturaleza y estructura del error, en general, y de su aplicación específica al caso del error moral, se siguen importantes consecuencias para el modo en que Sócrates considera el fenómeno del conflicto motivacional. Asimismo, se muestra la importancia decisiva que tuvo la concepción socrática para la discusión de la estructura de la conciencia moral tanto en Platón y Aristóteles como en Tomás de Aquino.

PALABRAS CLAVE: AUTOENGAÑO, CONCIENCIA MORAL, ERROR, RACIONALIDAD INTERNA

* El presente artículo es una versión revisada de una conferencia de apertura del año académico dictada en la Facultad de Filosofía, de la Universidad Panamericana, México D. F., el 31 de enero de 2012. Las secciones dedicadas a Sócrates, Platón y Aristóteles recogen desarrollos presentados de modo más detallado en trabajos precedentes (Vigo, 2002: 65-101; 2000: 509-543, reproducido en 2001: 5-41). La presente versión del texto fue redactada como parte del Proyecto de Investigación “Filosofía Moral y Ciencias Sociales” (FFI2009-09265) del Ministerio de Ciencia e Innovación (España).

** Departamento de Filosofía-Instituto Cultura y Sociedad (ICS), Universidad de Navarra, avigo@unav.es

THE ERRONEOUS CONSCIENCE. FROM SOCRATES TO THOMAS AQUINAS

Abstract: *In the realm of moral action, the socratic principle nobody errs willingly implies that the rational agent always chooses to do what he/she takes to be good or better for himself/herself: his/her choice is internally rational. Socrates' view of the connection between error and self-deception is a major contribution to western philosophy. The application of this view to the particular case of moral error has important consequences concerning the possibility of motivational conflict. Not only Plato and Aristotle but also Aquinas are strongly influenced by Socrates in their views concerning the structure of moral conscience.*

KEY WORDS: SELF-DECEPTION, MORAL CONSCIENCE, ERROR, INTERNAL RATIONALITY

I

Según el testimonio de Platón y Aristóteles, Sócrates defendió en el campo de la ética filosófica una concepción de apariencia altamente contraintuitiva. Su base se hallaría en el principio, cuya formulación más difundida, aunque no necesariamente más precisa, establece que *nadie hace el mal de manera voluntaria*. Así formulado, dicho principio presenta un carácter abiertamente contraintuitivo, porque de forma habitual damos por sentado que quienes obran de un modo que, desde el punto de vista de la evaluación moral, aparece ya como malo o reprobable, ya como bueno o elogiado, merecen reprobación o elogio, justo en la medida en que, en algún nivel de análisis, pueda decirse que escogen voluntariamente y deciden obrar tal como lo hacen. En este sentido, la voluntariedad y la intencionalidad parecen ser condiciones indispensables de la existencia misma de acciones plenas, vale decir, de acciones que puedan ser imputadas como tales al agente que las produce, y que, para bien o para mal, esto es, en mérito o en demérito, puedan ser tomadas como base para enjuiciar su comportamiento, desde el punto de vista de una evaluación específicamente moral. Sócrates cuestionó abiertamente esta suposición básica, y formuló, en cambio, la tesis que identifica el vicio o la maldad moral (*kakía*) con la ignorancia (*áгноia*), cuya contrapartida es la identificación de la virtud moral con el conocimiento o la ciencia (*epistéme*) (Aristóteles, EN: VI 13, 1144b 28-30).

La contraintuitividad de la posición socrática no puede verse como una simple apariencia superficial. Pero resulta agudizada, en buena medida, por el modo en que se formula de manera habitual el principio básico asumido por Sócrates en nuestra lengua y en otras lenguas modernas. Como es sabido, en la versión tradicional, derivada de la explicación platónica (*Protágoras* 345d; Aristóteles, EN: VII 3, 1145b 21-27), dicha formulación reza: *oudeis hekòn hamartánei*. En el griego de la época clásica, el verbo *hamartánein* no alude en primer lugar al error o la falta de índole específicamente moral, sino que tiene una significación más amplia de alcance predominantemente cognitivo. El verbo significa básicamente “errar”, “equivocarse”, y no específicamente “hacer el mal” o “pecar”. El sustantivo *hamartía*, relacionado de forma directa con dicho verbo, tampoco tiene habitualmente en el griego clásico la significación que presenta después en el *Nuevo Testamento*, donde se emplea en el sentido específico de “pecado”.¹

Atendiendo a la significación básica de la noción griega de *hamartía*, la fórmula socrática podría traducirse, de modo menos confuso, como “nadie yerra voluntariamente”. Traducido así, este principio pierde ya buena parte de su aspecto contraintuitivo, pues la noción de error posee también en nuestra lengua un carácter predominantemente cognitivo, y excluye, en su uso más frecuente, la voluntariedad del acto así descrito. Por otra parte, la versión cognitivista también permite entender me-

¹ Aunque ya en la época clásica admiten usos de alcance genuinamente moral, también los términos latinos *peccare* y *peccatum* (o la forma paralela *peccatio*) se emplean —y, tal vez, de manera predominante— con un significado básicamente cognitivo, que cubre también el mismo campo semántico que los términos *hamartánein* y *hamartía* en el griego clásico. En tal sentido, la versión latina que traduce el principio socrático como *nemo sua sponte peccat* no apunta todavía necesariamente en una dirección distinta de la formulación original griega. Por su parte, el adjetivo griego *hekón*, vertido de manera habitual por el adverbio “voluntariamente”, puede ser traducido de ese modo de forma legítima, pero sólo a condición de que se tome el término en un sentido lo suficientemente amplio y coloquial, como para que no se comprometa todavía con ninguna concepción determinada de la voluntad, ni con una determinada teoría de las facultades del alma y sus relaciones mutuas. De hecho, la expresión griega tiene también el sentido, mucho más modesto, que corresponde a giros españoles del tipo de “de buen grado”, “de buena gana”, entre otros.

por la conexión del principio de involuntariedad de la falta moral con la equiparación de la virtud con el conocimiento y el vicio (*kakía*) con la ignorancia que Sócrates parece haber llevado a cabo. Sin embargo, la posición socrática conserva todavía un núcleo en apariencia irreductible de contraintuitividad, que no obedece ya a la formulación misma del principio “nadie yerra voluntariamente”, sino, más bien, al hecho de que apele a dicho principio como fundamento de una concepción específica del error *moral*. En efecto, suponemos habitualmente que el error específicamente moral se distingue del cognitivo, es decir, del error a secas, justo por el hecho de que el primero implica, de uno u otro modo, la voluntariedad del acto así caracterizado.

Aplicado al ámbito de la acción moral, el principio socrático de que nadie yerra voluntariamente implica, como su conversa, la tesis según la cual toda vez que un agente elige algo, lo hace por considerarlo, al mismo tiempo, como bueno o, al menos, preferible a otra cosa. Esta tesis constituye un principio básico en la teoría de la acción subyacente no sólo a la ética socrática, sino también, de diversos modos, a la ética clásica de filiación platónico-aristotélica, en su conjunto. Dicho principio establece que el fin de una acción, en tanto objeto del deseo que motiva su producción, siempre es intencionado, al mismo tiempo, como un bien (*sub specie boni/sub ratione boni*). Se trata aquí de un principio de alcance descriptivo, y no normativo, lo que implica que su lugar sistemático ha de buscarse, propiamente, en el ámbito de la teoría de la acción, y no en el de la ética, como tal.² De este modo, es importante señalar que en la formulación del principio la noción de bien ha de entenderse en un sentido formal-funcional, y no estrictamente material, que cubre tanto el ámbito del bien real como el del aparente. Así entendido, el principio no pretende sugerir que todo agente de praxis escoja efectivamente como fin, en todas y cada una de sus acciones, aquello que constituye el *verdadero* bien. Tal sugerencia sería simplemente absurda, ya que declararía inexistente de hecho la falta *moral*, como tal.

² Para una defensa de la tesis de que el llamado “intelectualismo socrático” debe verse, más bien, como una posición situada en el nivel de consideración propio de la filosofía de la acción y no en el de la ética normativa, véase Gómez Lobo, 1998: 32 y ss.

Lo que dicho principio establece es, más bien, un postulado mínimo de *racionalidad interna* de las acciones, que constituye un presupuesto imprescindible para poder dar cuenta de ellas, en términos de la adscripción de determinados deseos y creencias al agente del caso. En efecto, explicar las acciones de los agentes racionales siempre implica, en algún nivel de análisis, *racionalizarlas internamente*. Ello es así, porque dar cuenta de las acciones, desde el punto de vista de su producción, no consiste, en definitiva, sino en presentarlas como motivadas por determinados deseos y creencias, que, sobre la base de la consideración de dichas acciones, pueden o incluso deben adscribirse al agente del caso. Pero esto significa, al mismo tiempo, considerar las acciones así explicadas, al menos en algún nivel de análisis, como congruentes con tales deseos y creencias del agente y, en tal sentido, como *internamente racionales*:³ sólo

³ Utilizo las nociones de “racionalidad interna” y “racionalización interna” en el sentido que tienen habitualmente en la teoría de la acción de autores contemporáneos como Donald Davidson. Tal sentido alude a la *consistencia interna* del conjunto de deseos y creencias de un agente, así como de las decisiones y acciones producidas sobre la base de dicho conjunto de deseos y creencias. En este sentido, una decisión o una acción es *internamente irracional*, cuando no resulta consistente con los deseos y creencias del propio agente acerca de lo que sería mejor para él mismo escoger, decidir y/o hacer. La noción de racionalidad interna expresa, pues, una exigencia débil o mínima de racionalidad, que, como tal, no toma en cuenta todavía por sí mismo el contenido proposicional de los deseos y creencias involucrados, salvo, justamente, en la medida necesaria para establecer si son o no consistentes entre sí. Al aspecto vinculado con la evaluación y el enjuiciamiento de las creencias en atención a su contenido proposicional específico apunta, en cambio, la noción de *racionalidad externa*, la cual presenta mayores requerimientos, ya que pone en juego, en la evaluación de los deseos y creencias de un agente, patrones de racionalidad dotados de pretensión de validez intersubjetiva. Así, por ejemplo, creer en la magia o en la astrología pasan por ser, al menos para los patrones de racionalidad característicos de la actual civilización técnico-científica, ejemplos de creencias *irracionales*, en el sentido de la noción fuerte de *racionalidad externa*, y ello porque tales creencias no parecen satisfacer adecuadamente las exigencias impuestas por dichos patrones de racionalidad. Sin embargo, un agente situado en el horizonte cultural de dicha civilización todavía puede creer de modo *internamente racional* en la magia o la astrología, en la medida en que dichas creencias puedan articularse de manera consistente dentro del conjunto de

logramos explicar adecuadamente la conducta de un agente, esto es, darle sentido, considerada desde la perspectiva de la tercera persona, cuando intentamos reconstruir de modo coherente los motivos que llevaron a dicho agente a producir las correspondientes acciones. Esta exigencia de racionalización interna, como sustento de la posibilidad de la explicación e interpretación de la conducta de un agente, vale incluso allí donde sus acciones puedan o deban ser consideradas como irracionales, *en sentido externo*. Por ejemplo, al reconstruir los motivos que condujeron o pudieron conducir a un imputado a cometer un crimen, el juez que entiende en la causa debe esforzarse por reconstruir de modo consistente un conjunto de deseos y creencias relevantes, atribuibles al imputado, a partir del cual la acción así realizada pueda verse como el resultado esperable del intento por traducir en concreto tales deseos y creencias. Pero, al proceder de este modo, el juez intenta aplicar al caso concreto del crimen, que configura en sí mismo una acción externamente irracional, el mismo principio básico de racionalización interna que asumimos en la explicación de *todas* las demás acciones —es decir, también de las externamente racionales—, con el fin de que puedan contar como acciones plenas, justo, en la medida en que reflejan en concreto los deseos y las creencias del agente que las produce de manera voluntaria.

En el plano de consideración correspondiente a la perspectiva de la primera persona, lo anterior puede formularse en términos que están directamente vinculados con los presupuestos básicos de la concepción en torno a la agencia racional que, cada uno de diferentes modos, asumen autores como Sócrates, Platón y Aristóteles. Los agentes racionales, en tanto lo son, se caracterizan por la capacidad de orientar su comportamiento de acuerdo con su propia decisión deliberada y con arreglo a

sus otros deseos y creencias. Sobre esta base, en el caso de tal agente, también las acciones motivadas por su creencia en la magia o en la astrología o, al menos, algunas de ellas (por ejemplo, gastar determinada cantidad de dinero en sortilegios o en horóscopos) podrán ser consideradas como *internamente racionales*, en la medida en que resulten congruentes con los deseos y creencias relevantes del propio agente. Para las nociones de racionalidad interna y externa en el sentido indicado véase Davidson, 1982: 289-305; reproducido en: 2004a: 169-187.

una cierta consideración global de la propia vida, por poco nítida y diferenciada que ésta pueda ser en algunos casos. Ella adquiere expresión en un conjunto de objetivos o fines de mediano y largo plazo. Por lo mismo, puede decirse que lo propio de un agente racional, en cuanto tal, es buscar a través de todas y cada una de sus acciones particulares aquello que juzga que lo beneficia (o lo beneficia más) o, al menos, que no lo perjudica (o lo perjudica menos); mientras que, inversamente, tiende a evitar todo aquello que juzga que lo perjudica (o lo perjudica más). Dicho juicio del agente acerca de lo que le resulta (más/menos) beneficioso o (más/menos) perjudicial puede siempre, por supuesto, resultar erróneo. Pero es imprescindible suponer su presencia, como fundamento de la toma de decisiones que un agente racional realiza en concreto, pues sería, sin más, inconsistente asumir que éste escoge voluntariamente lo que juzga menos benéfico o más perjudicial para sí mismo, al menos, en el mismo momento y en el mismo respecto en que lo escoge voluntariamente.

II

Sobre la base de lo dicho, se pueden comprender mejor las razones por las cuales Sócrates se ve inclinado a aplicar al caso específico del error moral su concepción de la naturaleza del error (cognitivo), en general. Según esta concepción, resulta esencial a *toda* forma de error la presencia de un componente irreductible de autoengaño. La tesis socrática sobre la conexión estructural entre error y autoengaño constituye, sin duda, uno de los aportes más decisivos de Sócrates al pensamiento filosófico occidental. Su significado nuclear puede formularse del siguiente modo: a quien está en el error, y en la medida precisa en que lo está, el error no se le revela como error, sino que se le aparece, más bien, como lo contrario de lo que en realidad es, es decir, como genuino saber. Precisamente por ello puede decirse que el sujeto se encuentra en el error.

Si, por caso, alguien cree erróneamente que la capital de Francia es Berlín, éste, en la medida en que lo cree, da por verdadera la proposición “la capital de Francia es Berlín”, la cual es, sin embargo, falsa. Pero dicho sujeto se encuentra en el error porque, mientras cree que la proposición

en cuestión es verdadera, no está en condiciones de reconocer la asimetría entre el valor de verdad que él mismo asigna a tal proposición a través de la correspondiente actitud proposicional (*verbi gratia*, la actitud de creencia, en el sentido de “creer que...”), por un lado, y el valor de verdad que corresponde a la proposición, como tal, por el otro. Según una caracterización tradicional, se está en el error cuando se tiene lo verdadero por falso o bien lo falso por verdadero. Pero esto sólo es posible cuando no se está en condiciones de reconocer, como tal, dicha divergencia en los valores de verdad vinculados, en cada caso, con las proposiciones y las correspondientes actitudes proposicionales. En este sentido, el error involucra esencialmente un componente de autoengaño de parte del sujeto que se encuentra en él, ya que mientras permanece en él, y en la medida en que lo está, el sujeto no está en condiciones de reconocer el error *como error*. Por el contrario, el sujeto que se encuentra inmerso en el error se ha identificado siempre ya con el correspondiente contenido proposicional, al darlo por verdadero, de modo tal que éste le aparece, más bien, como vehículo de saber. Tanto el error como el saber comportan esencialmente un componente de identificación, en virtud del cual el sujeto del caso se ha identificado siempre ya con los correspondientes contenidos proposicionales. Es justamente este rasgo compartido, es decir, la presencia en ambos de dicho componente de identificación, lo que explica que, *desde la perspectiva de la primera persona*, la diferencia entre error y saber tenga un carácter esencialmente auto-ocultante: *el error sólo puede operar efectivamente como error sustrayéndose en su carácter de error*.

En la medida en que involucra dicho componente de identificación de parte del sujeto y posee un carácter esencialmente auto-ocultante, el error sólo puede ser detectado al ser reconocido como tal, a través de un autodistanciamiento que quiebra la inicial identificación con los correspondientes contenidos proposicionales. Para retomar el ejemplo anterior, si alguien cree por error que la capital de Francia es Berlín, sólo podrá salir del error si reconoce como falsa la proposición “la capital de Francia es Berlín”, que inicialmente daba por verdadera, a través de la correspondiente actitud proposicional (*verbi gratia*, la creencia). Al hacer esto, el sujeto se distancia del contenido que antes tomaba como vehículo de saber y ahora, al declararlo falso, toma en retrospectiva como el

vehículo al error. Sólo entonces, es decir, una vez que el sujeto ha reconocido la falsedad de su creencia anterior el sujeto está en condiciones de obtener genuino saber: después de haber descubierto que, contra lo que creía saber, Berlín no es la capital de Francia, sino de Alemania, el sujeto puede reconocer que, en realidad, no sabe cuál es la capital de Francia y de buscar entonces los caminos para cubrir esa laguna en su conocimiento. Ahora está en mejores condiciones para llegar a saber en verdad cuál es la capital de Francia, precisamente, en la medida en que supera el auto-engañó inicial y toma conciencia de que no sabe lo que creía saber. Este simple ejemplo permite ilustrar con suficiente claridad el alcance de la tesis socrática según la cual el primer paso para la obtención de genuino saber reside habitualmente, al menos, en el caso de un ser falible y finito como el hombre, en el reconocimiento del error y la ignorancia como tales.

De esta concepción en torno a la naturaleza y la estructura del error, en general, y de su aplicación específica al caso del error moral, se siguen importantes consecuencias para el modo en que Sócrates considera diversos casos del fenómeno del conflicto motivacional. Por una parte, la posición general basada en la tesis según la cual nadie yerra voluntariamente, con la consiguiente puesta de relieve del componente de autoengaño involucrado necesariamente en toda forma de error *moral*, en la medida en que, como *error*, involucra una dimensión esencialmente cognitiva, trae consigo, a modo de corolario, la famosa tesis socrática consistente en negar la posibilidad de la existencia de la acción incontinente, es decir, de aquella acción en la cual el agente racional escogería, a sabiendas, aquello que al mismo tiempo considera que es peor para él mismo.⁴

La incontinencia constituye uno de los principales casos de conflicto motivacional que dan cuenta de la posibilidad de la irracionalidad interna de las acciones, de modo tal que se encuentra, ya en virtud de su propia caracterización, en una incómoda relación de tensión con el principio metódico-hermenéutico que establece que explicar las acciones consiste, de uno u otro modo, en racionalizarlas internamente. Según una caracterización formal que goza, al menos *prima facie*, de una acep-

⁴ Para la adscripción a Sócrates de la tesis de la imposibilidad de la incontinencia véase Platón, *Protágoras* 352a-353b; Aristóteles, en VII 3, 1145b 22-24.

tación suficientemente amplia, una acción 'x' contaría como incontinente, si y sólo si su producción satisface las siguientes tres condiciones: a) el agente realiza la acción 'x' de modo intencional; b) hay una acción alternativa 'y' que el agente podría realizar; y c) el agente cree que, consideradas todas las cosas, sería mejor para él hacer 'y' que hacer 'x'.⁵ Para ser incontinente, una acción debe, por lo tanto, contar como a) intencional, b) no forzada ni inevitable, al menos desde el punto de vista de las creencias del propio agente, y además c) inconsistente con las creencias de dicho agente respecto de lo que sería bueno o mejor para él mismo. El núcleo del problema de la incontinencia se captura cuando se atiende, sobre todo, a la conjunción de a) y c), pues ella expresa la dificultad básica que los fenómenos de irracionalidad interna plantean, frente al requerimiento de racionalización que trae consigo todo intento de explicación de una acción, en su carácter de acción plena o intencional.

Sócrates parece haber sido el primer filósofo que identificó de modo preciso la tensión estructural aquí subyacente, a saber: la tensión entre el carácter intencional, por un lado, y el carácter internamente irracional, por el otro, de una y la misma acción. Y si la presentación de su posición que realizan Platón y Aristóteles es adecuada, habrá que decir que Sócrates resolvió la dificultad así detectada del modo más radical y coherente, pero también más contraintuitivo posible, al enfatizar la imposibilidad de conectar en una descripción consistente los requerimientos expresados por a) y c), y recalcar, con ello, la necesidad de apartarse del modo habitual de hablar, con el fin de redescubrir los supuestos casos de acciones incontinentes en términos de acciones basadas, de uno u otro modo, en ignorancia.

⁵ Para esta caracterización de la acción incontinente, véase Davidson, 1980: 21-42, en especial 22. Para una amplia discusión sistemática de la incontinencia dentro del ámbito más amplio de los fenómenos de irracionalidad interna, véase Mele, 1992. Una reconstrucción de las principales posiciones respecto de la incontinencia, desde Sócrates a Davidson, pasando por Platón, Aristóteles, Tomás de Aquino y Hare, se encuentra en Spitzley, 1992. Una buena discusión de las principales concepciones acerca del conflicto motivacional en la filosofía griega, desde Sócrates hasta los estoicos, se encuentra en Price, 1995.

El famoso pasaje del arte de la medida del *Protágoras* platónico sugiere que Sócrates pudo haber considerado, básicamente, la idea de que en los casos de acciones supuestamente incontinentes lo que tiene lugar es, en rigor, un cierto tipo de error cognitivo, que se produce al ponderar y comparar los bienes y los males o, más específicamente, los placeres y los dolores, involucrados en la decisión en favor o en contra de una determinada acción. Quien, para decirlo según el modo habitual de hablar, es subyugado por un placer, que se le presenta como cercano e inmediato, cae víctima de una suerte de ilusión óptica, parecida a la que tiene lugar en la percepción de los objetos en el espacio: los que están más cerca parecen más grandes que los lejanos, aunque éstos puedan ser mucho mayores. Sin embargo, en el caso del espacio, el arte de la medida permite corregir fácilmente esa apariencia. Del mismo modo, en el caso de la ponderación relativa de bienes (placeres) y males (dolores) presentes y futuros sería necesario también un correspondiente arte de la medida que, en la comparación de las diferentes intensidades, permita corregir el efecto distorsivo producido por la mayor o menor cercanía temporal, y restablecer así el adecuado orden de prioridades, de modo tal que el agente no escoja *ahora* un placer menor que *posteriormente* traiga un dolor mayor o, viceversa, no evite *ahora* un dolor menor que dé como fruto placeres mayores *en el futuro* (cfr., Platón, *Protágoras*: 353c-359a). La eficacia del arte de la medida residiría aquí, por lo tanto, en su poder de contrarrestar la fuerza de la apariencia (*he toû phainoménu dýnamis*), que induce al error en la ponderación de los bienes y males puestos en juego por la decisión del caso (cfr., Platón, *Protágoras*: 356d 4).

Ahora bien, si éste es el caso con las acciones pretendidamente incontinentes, no se está aquí, en realidad, en presencia de un genuino conflicto motivacional, del tipo que habitualmente se supone que les subyacería. Más que un conflicto de deseos opuestos operantes de manera simultánea en una y la misma situación de acción, lo que se tiene en estos casos es, según Sócrates, un único deseo básico, obtener el mayor bien posible, que sobre la base de un determinado error cognitivo, fundado en una incorrecta ponderación de los bienes concretos que están en juego, da lugar, sin embargo, a una acción que, *desde el punto de vista de la perspectiva externa propia de la tercera persona*, resulta inadecuada, pues no realiza de modo efectivo dicho deseo básico: queriendo obtener el mayor

bien, el agente termina escogiendo y, en definitiva, obteniendo un bien menor o incluso un mal. El paso de la perspectiva de la primera a la segunda o tercera persona resulta decisivo en este análisis, pues parece esencial a la posición socrática la asunción de que, al menos, en el momento mismo de la acción, el agente que escoge algo que resulta ser peor para él mismo, si en realidad lo escoge, debe considerar, en algún nivel de análisis, como mejor o preferible eso mismo que efectivamente escoge. De lo contrario, no habría modo de explicar por qué el agente escoge lo que escoge, lo que equivale a cuestionar, al mismo tiempo, la posibilidad misma de afirmar que lo haya escogido.

La experiencia muestra, por cierto, que los agentes muchas veces describen y lamentan *retrospectivamente* sus acciones, como si no hubiesen sido realizadas sobre la base del mejor juicio acerca de lo que resulta mejor para ellos mismos. Contamos en el lenguaje prefilosófico con una cantidad de recursos para dar expresión a actitudes en las que sometemos a crítica nuestras propias decisiones y nos arrepentimos de ellas o de sus consecuencias. Pero esto no provee una objeción decisiva contra la posición socrática, e incluso puede servir para prestarle una confirmación indirecta, si se considera el hecho de que dichas actitudes involucran un componente esencialmente reflexivo, que presupone la capacidad del agente de tomar distancia respecto de los propios motivos de acción, considerándolos en retrospectiva desde un punto de vista nuevo y diferente, que, como tal, no se deja reducir sin residuo al punto de vista desde el cual surgió originariamente la decisión ahora criticada. En el caso de la autocrítica, el lamento sobre el propio comportamiento y el arrepentimiento se trata, en efecto, de actitudes en las cuales el agente racional vuelve reflexivamente sobre sus propias decisiones y acciones, desde una perspectiva que, en lo que concierne al aspecto de falta de identificación con los motivos del obrar, resulta comparable a la que caracteriza a la crítica ejercida en segunda o tercera persona.

Así delineada, la posición socrática aparece constituida por un peculiar y complejo engarce de elementos, entre los cuales resultan fundamentales los siguientes. Por una parte, el principio de racionalización interna, que adquiere expresión en el *dictum* “nadie yerra voluntariamente”, y el reconocimiento de la presencia de un componente irreducible de autoengaño, como esencial a toda forma de error, conducen a

Sócrates al rechazo de la posibilidad de la existencia de fenómenos de irracionalidad interna, en el sentido fuerte que implica un conflicto motivacional verificado *en el plano sincrónico*, como sería el caso de la incontinencia, en el sentido más habitual del término. Pero, por otra parte, Sócrates no excluye la existencia, *en el plano diacrónico* de consideración, de fenómenos de autodistanciamiento, a través de los cuales el agente toma posición, de modo reflexivo y crítico, respecto de sus propios motivos de acción, las decisiones tomadas a partir de ellos y las consecuencias resultantes. Por el contrario, Sócrates concede a dichos fenómenos una importancia decisiva, cuando se trata de evaluar el papel que desempeñan con respecto a la posibilidad del progreso moral del agente. Mi tesis en este punto es que tanto el aspecto limitativo de la posición socrática, referido a la imposibilidad de la existencia de conflicto motivacional en el plano sincrónico, como el aspecto positivo, referido a la función insustituible que cumplen los fenómenos de autodistanciamiento para hacer posible el progreso moral de los agentes racionales, constituyen, en rigor, las dos caras de una misma moneda, en la medida en que ambos se fundan en la concepción socrática del error como un fenómeno epistémico que involucra esencialmente un componente de autoengaño.

Si nos atenemos al testimonio que ofrece Platón en el *Gorgias*, Sócrates habría concedido especial atención a dos dispositivos básicos que, a través de la producción de experiencias de contraste, ponen al sujeto autoengañado en condiciones de hacer el tipo de experiencia de autodistanciamiento que le ofrece la posibilidad de superar dicha situación, a saber: la refutación, en el contexto dialógico, y el castigo, en aquellos contextos, sean de tipo pedagógico o judicial, en los cuales los instrumentos del diálogo ya no bastan por sí solos para obtener el resultado esperado (*cfr.*, Platón, *Gorgias*: 505e-506c y 466a-474b, respectivamente). La refutación y el castigo constituyen factores de contraste que nadie escogería por sí mismos. En el ámbito teórico-cognitivo, el bien buscado por sí mismo es la verdad, así como en el ámbito práctico-moral lo es la justicia o, lo que es lo mismo, el bien. Sin embargo, en la hipótesis de que el sujeto se encuentre en el error respecto de la verdad sobre una determinada materia relevante de discusión, o bien respecto de su propio bien como agente racional, dichos factores de contraste adquieren una im-

portancia decisiva, en la medida en que ponen al sujeto del caso en condiciones de hacer las correspondientes experiencias de autodistanciamiento, a través de las cuales puede reconocer, por primera vez, su error *como error* y, así, escapar de la trampa del autoengaño.

En la posibilidad de hacer tales experiencias de autodistanciamiento juega, pues, un papel decisivo el factor irreductible de alteridad encarnado, respectivamente, por el interlocutor que comparte el diálogo crítico y por la instancia competente para administrar el correspondiente castigo. La propia naturaleza del error, en definitiva, hace necesaria la mediación de factores de alteridad que pongan a prueba la consistencia de las creencias en las que, como sujetos de conocimiento y como agentes morales, estamos siempre ya instalados, de modo tal de poner de manifiesto y, con ello, de desactivar en su eficacia ocultante y autoocultante, aquellas creencias en las cuales la identificación con determinados contenidos adquiere la forma de una conciencia errónea, fundada en el autoengaño respecto de la verdad o del bien.

Un aspecto ulterior, vinculado con la función positiva que Sócrates adscribe a los mencionados fenómenos de autodistanciamiento, queda, por el contrario, más bien, relegado al trasfondo, aunque su importancia sistemática, y su decisiva influencia para el posterior desarrollo de las concepciones medievales y modernas de la conciencia y la conciencia errónea, difícilmente podría ser exagerada. Se trata de la constatación elemental de la necesidad de la presencia de determinadas condiciones subjetivas que hagan posible la eficacia de aquellos dispositivos que, como la refutación o el castigo, apuntan a hacer posible el autodistanciamiento, la superación de la identificación en el modo de la conciencia errónea y, con ello, el acceso a nuevas formas, no erróneas, de la conciencia de sí y la identificación consigo mismo. En efecto, como muestra de modo dramático el propio texto del *Gorgias*, allí donde tales condiciones subjetivas no están dadas o bien han quedado desactivadas en su operatividad, ningún fenómeno de contraste inducido desde el exterior puede forzar, sin más, la apertura del sujeto autoengañado a una nueva comprensión de sí mismo y de su propio bien real.

Los casos de Polo y, sobre todo, de Calicles manifiestan, con admirable nitidez, algo que parece ser una lección elemental de la mayoría de los diálogos socráticos de Platón, a saber: que nadie puede ser conducido,

sin más, a la transparencia respecto de sí mismo, si él mismo no está dispuesto ya, en alguna medida, al riesgo de hacerse cargo de sí, precisamente, del modo en que tal exigencia de autotransparencia prescribe. Hay sin duda un núcleo de indelegabilidad individual en toda forma, recta o errónea, autotransparente o auto-ocultante, de hacerse cargo de sí mismo. Pero, además, hay que suponer también, al menos como posibilidad, cierta capacidad de autotransparencia en todos los agentes racionales, precisamente, en cuanto racionales, porque, de lo contrario, la instalación en las diversas formas de la conciencia errónea no podría ser superada, siquiera en parte, ya que, al menos, en la perspectiva de la primera persona, tampoco podría ser detectada, como tal.

III

No resulta necesario enfatizar la decisiva importancia que posee la posición socrática, cuando se trata de comprender, en algunos de sus puntos de partida básicos, las concepciones de los pensadores griegos posteriores, desde Platón y Aristóteles hasta los estoicos, en el ámbito de la filosofía de la acción y de la ética. La detección de diferentes aspectos en los cuales se haría sentir, de diversos modos, el influjo socrático ha jugado un papel central en la investigación especializada, tanto cuando se apunta a marcar la continuidad de determinados motivos o asunciones, como cuando se trata de poner de relieve la existencia de momentos de divergencia y toma de posición crítica frente al socratismo. Particularmente, en lo que concierne a la debatida cuestión del así llamado *intelectualismo socrático* y, sobre todo, en el caso de Platón y Aristóteles, no ha sido infrecuente que los intérpretes llegaran a un diagnóstico que combina ambos aspectos, marcando la presencia tanto de momentos de continuidad como de distanciamiento respecto de la posición de Sócrates. Desde diferentes perspectivas, diversos intérpretes han creído tener que concluir que, pese a introducir sustanciales modificaciones y refinamientos, sobre todo en lo que concierne al modo de dar cuenta de la producción de la acción desde el punto de vista psicológico, tanto Platón como Aristóteles conservaron, en definitiva, un núcleo socrático irreductible, en su modo de concebir la acción racional, tanto desde el punto de vista motivacio-

nal como en atención a su calificación estrictamente moral. Ambos retoman y, en cierto modo, radicalizan el motivo socrático, según el cual el camino del progreso moral, al igual que el de la superación del error y la ignorancia, transita necesariamente por las experiencias de autodistanciamiento que facilitan determinados factores de contraste. Una breve enumeración de puntos centrales, que he discutido más ampliamente en otro lugar (Vigo, 2002) puede servir para ilustrar lo dicho.

1) A través de una nueva concepción de la motivación de la acción, basada en una concepción psicológica mucho más diferenciada, Platón y Aristóteles dan cuenta de la posibilidad de la existencia de genuino conflicto motivacional, que, como tal, no queda limitado al plano diacrónico, sino que se extiende también al *plano sincrónico* de consideración.

2) Por otro lado, Platón y Aristóteles extienden la temática socrática del autodistanciamiento posibilitado a través de factores de contraste, más allá del ámbito de las creencias, también al ámbito de *los deseos* del agente racional. Todo ello les permite enfatizar, de un modo nuevo y diferente, las funciones positivas que poseen los fenómenos, en sí mismos no deseables, de división interior, allí donde se trata de dar cuenta de la posibilidad de que el agente deje atrás aquellas formas perversas de la conciencia de sí, que están basadas en el autoengaño respecto del propio bien. Por medio de una concepción diferente de la naturaleza de la motivación psicológica, que apunta a hacer justicia a la complejidad de factores intervinientes en la producción efectiva de las acciones por parte de los agentes racionales, tanto Platón como Aristóteles trasladan el modelo socrático, que identifica en los factores de contraste la instancia de mediación y alteridad que hace posible el autodistanciamiento y la vuelta reflexiva sobre sí, al interior mismo del alma individual del agente. Se trata, en cierto sentido, de una interiorización del modelo socrático que, sin abandonar la lección referida a la necesidad de la mediación de factores externos de alteridad y contraste, apunta a proyectar dicha estructura de mediación a la propia realidad interior del agente racional. De este modo, Platón y Aristóteles superan lo que podría verse como los aspectos más débiles o poco diferenciados de la posición socrática, que conciernen, sobre todo, al modo en que Sócrates tiende a comprender

el proceso de la motivación y la producción de la acción, sobre la base de la asunción de una suerte de monismo psicológico, que no permite hacer debida justicia a los fenómenos de división interior y conflicto motivacional.

3) Finalmente, también el modo en que Sócrates piensa la conexión entre el error moral y, en general, el vicio, por un lado, y la ignorancia, por el otro, queda sujeto a una revisión decidida que permite introducir una considerable diferenciación adicional. En particular, Aristóteles enfatiza el hecho de que la conexión entre vicio e ignorancia no puede esclarecerse adecuadamente si no se lleva a cabo una distinción de los diferentes tipos de ignorancia que pueden intervenir en la motivación y la producción de las acciones. Aristóteles introduce aquí dos distinciones cruciales que permiten hacer justicia a una cantidad de aspectos vinculados con la imputación de las acciones y con la evaluación de su calidad moral, a saber: por un lado, la distinción entre la ignorancia universal referida a los principios normativos a los que deben quedar sujetas las acciones y la ignorancia particular referida a las circunstancias y las marcas situacionales de la acción; por otro, la distinción entre la ignorancia como causa de la acción y la ignorancia como estado en el que la acción se lleva a cabo.⁶ De este modo, Aristóteles elabora una teoría altamente diferenciada, que permite dar una respuesta convincente a preguntas que la posición socrática parecía dejar abiertas o precisadas de manera insuficiente, tales como, la pregunta por las precisas condiciones bajo las cuales la ignorancia del agente puede o debe contar como fundamento de la involuntariedad de la acción y como causal de eximición de responsabilidad.

Ahora bien, sería un error pensar que a través de todas estas modificaciones y precisiones Platón y Aristóteles dejan atrás para siempre la posición alcanzada por Sócrates. En efecto, lejos de pretender eliminar todo residuo de socratismo de sus propias concepciones, Platón y Aristóteles parecen apuntar, más bien, a preservar la intuición nuclear referida al componente de autoengaño presente en todo error y a la estructura

⁶ Cfr. Aristóteles, *EN*: III 2, 1110b 24-27, donde se introduce la distinción entre obrar *en ignorancia* (*agnoôn*) y *por ignorancia* (*di' ágnouan*).

de la conciencia errónea, pero recontextualizándola del modo requerido, con el fin de no asumir algunas de las consecuencias más paradójicas de la posición socrática, tales como la identificación del vicio con la ignorancia, sin otras especificaciones, o la negación de la posibilidad del conflicto motivacional y la incontinencia. Que la intuición socrática nuclear queda aquí finalmente preservada, se advierte de inmediato, cuando se considera el hecho de que tanto Platón como Aristóteles se mantienen firmes en la convicción de que el único modo posible de explicar la producción de las acciones, desde la perspectiva de la primera persona, consiste en asumir que el fin de la acción es siempre intencionado bajo la especie o figura del bien. Ello implica que resulta imposible suponer que cuando el agente racional escoge voluntariamente un determinado curso de acción encaminado a lograr un determinado objetivo podría estar escogiendo aquello que considera, al mismo tiempo, malo o peor que otra cosa, al menos, en el mismo momento en que lo escoge y en el aspecto bajo el cual lo escoge.

IV

Por último, quisiera considerar brevemente el modo en el que la intuición socrática nuclear referida al componente de autoengaño —presente en todo error y la estructura de la conciencia errónea— se proyecta también en la concepción de Tomás de Aquino. Como es sabido, éste elabora su propia teoría de la motivación y la producción de las acciones en estrecha conexión con la concepción aristotélica, pero también apela a una cantidad de otras fuentes y autoridades, que, además del pensamiento propiamente cristiano, en muchos casos están vinculadas también a la tradición del (medio- y neo-) platonismo y del estoicismo.

No trataré la posición de Tomás de Aquino a partir de lo que pudiera parecer la vía natural de acceso para el problema que nos ocupa, la cual sería su tratamiento de la acción incontinente. Aquí Tomás sigue las líneas generales de la discusión aristotélica, más allá de claras diferencias de matiz y acentuación en determinados puntos de detalle. Por lo mismo, su posición podría caracterizarse con facilidad en términos básicamente coincidentes con la de Aristóteles, a saber: como un intento de

explicar la posibilidad de la acción incontinente, sobre la base de una teoría de la motivación que da lugar a los fenómenos de división interior y conflicto motivacional, pero reteniendo, a la vez, la idea de que el incontinente elige actuar como actúa, porque opta por lo que le parece mejor, en el momento mismo de la elección y en el respecto en que lo elige; ello muy a pesar del hecho de que tal juicio deba verse, en definitiva, como el emergente de una desactivación transitoria de sus propias convicciones más permanentes acerca de lo que sería mejor para él mismo, producida por la influencia perturbadora de factores no-cognitivos (emociones, deseos apetitivos) sobre la capacidad de evaluación de la situación concreta de acción.

Donald Davidson que en su discusión del fenómeno de la incontinencia toma a Aristóteles y Tomás como ejemplos paradigmáticos de este tipo de explicación, los critica por no superar de modo definitivo el punto de partida socrático, lo que los llevaría a la imposibilidad de considerar lo que él denomina *incontinencia con mente despejada* o, tal vez podríamos decir, *a sangre fría*, que sería el fenómeno que representa el genuino conflicto motivacional en el plano estrictamente sincrónico.⁷

A mi modo de ver, el propio Davidson sobreestima, en alguna medida, las diferencias reales entre su posición y las concepciones clásicas en la línea de Platón, Aristóteles y Tomás, quizá porque no advierte del todo que el desafío planteado por Sócrates se puede dirigir también contra su propia posición.⁸ Pero no me internaré aquí por este camino. Me limitaré, más bien, a un par de observaciones referidas al modo en el que Tomás intenta responder a la pregunta de si la conciencia errónea es o no moralmente vinculante en el notable texto de *STh* I-II q. 19 a. 5. En la respuesta a esta cuestión se patentiza de un modo particularmente nítido

⁷ En este sentido, véase la distinción entre formas sincrónicas y diacrónicas de irracionalidad interna en Davidson, 1985: 345-354; reproducido en Davidson, 2004b: 189-198. Véase en esp. p. 353 de la versión original (p. 197 de la reproducción). Thomas Spitzley considera como uno de los principales méritos de la posición de Davidson el haber (supuestamente) dado cuenta por primera vez de los fenómenos de irracionalidad interna estrictamente sincrónica, tales como la incontinencia a sangre fría. Véase Spitzley, 1992: 215-223.

⁸ Para este punto, remito a la discusión en Vigo, 1998: 51-97, en esp. 76 y ss.

do de qué manera la concepción tomasiana incorpora y asimila de modo creativo los aspectos medulares de la concepción socrática del error. Señalo, de modo muy escueto, los puntos principales que permiten caracterizar la posición fijada por Tomás.

1) El contexto de la discusión sobre el carácter vinculante (o no) de la conciencia errónea es una consideración mucho más amplia acerca de lo que puede llamarse el aspecto interno de la moralidad de los actos, es decir, el aspecto referido a las condiciones subjetivas de la moralidad de las acciones. Siendo una tesis básica de la posición tomasiana que la moralidad de los actos viene definida primariamente por un conjunto de requerimientos externos u objetivos, como son la referencia al objeto del acto y la correspondencia con el patrón normativo provisto por la Ley Eterna, no menos cierto es que Tomás asume, de manera explícita, la tesis de que tales condiciones o requerimientos externos son condiciones necesarias, pero no suficientes, para la moralidad de la acción. Ésta pierde su bondad no sólo allí donde colisiona con las exigencias que derivan, inmediata o mediatamente, de la Ley Eterna, sino que también la pierde o puede perderla allí donde, siendo conforme exteriormente con tales exigencias, no reúne el conjunto de condiciones interiores o subjetivas que dan cuenta del hecho de que lo querido en cada caso no sólo es lo que debe quererse, sino que es querido, además, del modo en que debe quererse, esto es, por sí mismo y sobre la base del correspondiente conocimiento, y no de modo puramente instrumental o por accidente.

2) Por lo mismo, y teniendo en cuenta que, por otra parte, el acceso a la Ley Eterna siempre está mediado por la razón, sea a través de la presencia en ella de la llamada Ley Natural, o bien mediante el reconocimiento racional de la autoridad de la Escritura, Tomás sostiene que la bondad de los actos de la voluntad depende *en definitiva* de su conformidad con la Ley Eterna, pero que su causa próxima o inmediata es, más bien, la conformidad con la razón, que es la vía a través de la cual nos viene dada *de modo natural* la Ley Eterna, y la única disponible allí donde no estuviera dada o accesible una revelación de tipo sobrenatural (*cfr.*, *STh* I-II q. 19 a. 4).

3) Sobre esta base, la pregunta clave es si una razón o conciencia errónea es o no vinculante (*utrum ratio errans obliget*), aceptando, por otra parte, que el error de la razón o conciencia, cuya corrección deriva en último término de la conformidad con la Ley Eterna, no puede provenir él mismo de esa Ley Eterna, sino que debe tener otras causas, que pueden ser múltiples. La respuesta de Tomás no puede ser más clara: la conciencia es siempre vinculante, o con mayor precisión, *subjetivamente vinculante*, también allí donde es o pueda ser errónea. Frente a esto, no faltará quien pudiera sorprenderse y preguntarse cómo puede un autor objetivista en moral, como Tomás, mantener la postura indicada. Y aunque no faltara tampoco quien pudiera tener la tentación de ver aquí problemas intrincados o, incluso, imposibles de resolver, la respuesta resulta, a mi modo de ver, poco menos que elemental: Tomás es objetivista en materias de fundamentación (última) de la moralidad, pero no es externalista en su concepción relativa a la constitución del valor moral de las acciones humanas, ya que considera igualmente necesarios tanto factores externos u objetivos como factores internos o subjetivos para dar cuenta de la bondad *moral* de la acción, la cual no se agota ni podría agotarse en su carácter externamente benéfico e incluso puede, en ocasiones, estar reñida con él.

4) Por consiguiente, Tomás rechaza en el texto los intentos de los teólogos morales que buscan reducir el problema relativo al carácter vinculante de la conciencia a través de consideraciones de corte externalista, referidas al carácter bueno, malo o neutral del acto, por referencia meramente a su objeto. Estas consideraciones conducen a la tesis según la cual, allí donde la razón o la conciencia es errónea, sería bueno obrar en contra de la propia razón o conciencia, porque, en tal caso, el acto será objetivamente bueno.⁹ Tomás considera esta posición, de modo

⁹ A este problema apunta también el argumento sofístico referido por Aristóteles en *EN*: VII 3, 1146a 21-31, que puede resumirse en la fórmula *aphrosýne* + *akrasía* = *areté*. Es decir: alguien que, 1) en virtud de su insensatez, quisiera voluntariamente obrar mal, pero, al mismo tiempo, 2) fuera incontinente y obrara así en contra de sus propias convicciones acerca de lo mejor para él mismo, produciría finalmente acciones conformes

explícito, como *no razonable (irrationabiliter dicitur)*, precisamente, porque no toma debidamente en cuenta las condiciones internas o subjetivas del acto. No es que, en caso de ir en contra de su mejor juicio, el agente meramente abandona lo que le indica *su razón*. A ello se añade que, al obrar así, eventualmente se obligará a perseguir aquello que a él mismo se le aparece incluso como malo, al menos, cuando se trata de lo contrario de lo que se le aparece como bueno, pero, en cualquier caso, siempre será algo que se le aparece como menos bueno y, por lo tanto, como menos preferible a aquello que su mejor juicio le señala como más deseable. La tesis según la cual sería bueno obrar en contra de la propia razón o la propia conciencia, allí donde ésta es errónea, nos pondría, pues, frente a lo que podría denominarse *incontinente recomendada* o, si se prefiere, *irracionalidad interna recomendada*. En efecto, se estaría recomendando al agente que sea, sin más, *internamente irracional* en sus comportamientos.

5) Como es obvio, Tomás no puede admitir semejante cosa: todo lo que el agente persigue y debe perseguir es algo que, en la medida en que lo hace, se le aparece y debe aparecersele como bueno. El error en el cual incurre la razón, explica Tomás, no procede de Dios, como sí proceden de Él la razón y la Ley Eterna presentes en ella. Sin embargo, el agente sumido en el error persigue aquello que no es *realmente* bueno (o mejor) siempre bajo la especie del bien, esto es, considerándolo bueno (o mejor), cabe decir: persigue el mero bien aparente, tomándolo, sin embargo, como un bien real. Y en esa misma medida, y si es verdad que toda verdad procede, en último término, de Dios, su juicio erróneo lo lleva a perseguir, en cierto modo, como procedente de Dios justamente aquello que no podría proceder de Él: la persecución del bien aparente es, pues, ella misma, en definitiva, un cierto modo de perseguir el bien, precisa-

a la virtud. Como es obvio, Aristóteles no puede aceptar la conclusión, por la sencilla razón de que la virtud ética no puede ser vista, a su juicio, como un resultado accidental de la concurrencia de dos disposiciones viciosas. Pero el argumento es ilustrativo del tipo peculiar de fenómeno que, de modo no demasiado afortunado, se denomina a veces “incontinencia (*akrasía*) inversa”.

mente porque el bien aparente es perseguido siempre, incluso bajo condiciones de autoengaño, como un bien real, y no como uno aparente.¹⁰

6) Por cierto, Tomás concede a San Agustín, a quien cita en el texto, que la orden de una autoridad inferior no puede contradecir ni anular la orden de una superior. Pero, precisando el alcance de este principio, y en una vena característicamente socrática, señala también que el propio reconocimiento de la verdad de dicho principio excluye ya, como tal, la posibilidad de una *completa* sumisión en el error: no puede engañarse *completamente* quien está en condiciones de acceder a la verdad del principio que le permite deponer su propio juicio respecto de una situación particular, para acudir a la Palabra de Quien, a diferencia de él mismo, ya no podría errar. Aparece aquí una vez más, como se puede ver, la idea de cuño socrático-platónico según la cual una genuina superación del autoengaño jamás puede ser motivada ni mucho menos provocada de un modo puramente exterior, sino que su misma posibilidad presupone ya la capacidad del sujeto de distanciarse de sí mismo y reconocer así su error como error.

V

Si la discusión desarrollada hasta aquí resulta en alguna medida convincente, habrá que reconocer la decisiva importancia que posee la concepción socrática para dar cuenta del origen de la reflexión filosófica griega y, en general, clásica, en el campo de la ética, y del rumbo que dicha reflexión tomó en algunos de los representantes más importantes de dicha tradición de pensamiento: Platón, Aristóteles, Tomás de Aquino.

¹⁰ Naturalmente, esto vale también allí donde lo que se persigue como un bien es la propia apariencia. Siguiendo la línea argumental antes esbozada habrá que decir en este caso que el agente que obra de ese modo lo hace sobre la base de la consideración según la cual la apariencia o cierto tipo de apariencia debe ser vista, al menos en determinados contextos o bajo determinadas condiciones, como algo *realmente* bueno o mejor que otra cosa.

En particular, en lo que concierne a la concepción socrática acerca de la naturaleza del error moral, en cuanto caracterizado esencialmente por involucrar un componente irreductible de autoengaño del agente respecto de su propio bien real, hay dos aspectos centrales, cuya influencia es notoria y decisiva. Por un lado, se trata de la intuición básica, subyacente a la posición de Sócrates, según la cual, a pesar de lo que podría parecer a primera vista, no hay, en definitiva, un modo realmente satisfactorio de dar cuenta de la posibilidad de la elección del mal por parte de los agentes racionales, si no se hace referencia a la presencia, en algún nivel de análisis, de alguna forma —total o parcial, transitoria o permanente— de ofuscación de la conciencia, en virtud de la cual el agente ya no acierta a reconocer con eficacia el genuino carácter de aquello que escoge para sí, al considerarlo, en algún sentido o respecto, como bueno y deseable cuando, en realidad, constituye un mal que lo daña en mucho mayor medida de lo que podría beneficiarlo. La influencia de este motivo central en la posición de Sócrates se manifiesta en el hecho de que tampoco Platón, Aristóteles y Tomás de Aquino parecen conceder la posibilidad de que un agente racional escogiera el mal, reconociéndolo a la vez, y con completa claridad de conciencia, como un mal, en el mismo momento y en el mismo respecto en que lo escoge.

Autores modernos como Davidson han querido ver en esto una deficiencia de las concepciones clásicas en torno a fenómenos de irracionalidad interna como la incontinencia, en la medida en que no dejarían espacio para la llamada incontinencia *a sangre fría* o *con mente despejada*, que representaría el caso más genuino de conflicto motivacional en el plano estrictamente sincrónico. Dejemos de lado aquí el hecho de que es argüible que el propio Davidson sobreestimó, en alguna medida, las diferencias reales entre su propia posición y las concepciones clásicas en la línea de Platón, Aristóteles y Tomás de Aquino. Sea como fuere, la pregunta de fondo es si realmente puede hacerse plausible la idea de que el agente racional escoge, en el momento de la decisión concreta de acción, algo que, en el mismo momento y en el mismo sentido en que lo escoge, considera a la vez como malo o peor para él mismo. ¿No queda, en definitiva, siempre abierta la sospecha de que dicha decisión, si ha de poder ser explicada de alguna manera, tendrá que poder ser redescrita, de uno u otro modo, en términos de la opción por algo que, desde la

perspectiva interna del agente en la situación particular en que la decisión es ejecutada, es considerado al mismo tiempo como bueno o preferible, en algún sentido o aspecto determinado, y escogido bajo la descripción que corresponde precisamente a dicho sentido o aspecto? ¿No reside aquí un elemento constitutivo de lo que la posterior tradición de la teología moral tematizó bajo el nombre de *mysterium iniquitatis*, cuyo carácter de misterio irreductible viene dado justamente por el hecho de que la opción por el mal parece sustraerse, por principio, a la posibilidad de ser explicada en la perspectiva interna de la primera persona en los mismos términos en que puede y debe ser explicada desde la perspectiva externa de la segunda o la tercera persona?

El carácter ubicuo de la posición elaborada por Sócrates en torno a la naturaleza del error, en general, y el error moral, en particular, como fenómenos que involucran necesariamente un componente irreductible de autoengaño, se revela ya en el hecho de que, más allá de las grandes diferencias en lo que concierne a la explicación de la producción de la acción y la posibilidad del conflicto motivacional, Platón, Aristóteles y Tomás de Aquino no creyeron posible eliminar de raíz todo vestigio de socratismo en sus propias concepciones.¹¹

No parece casual, pues, que algunos testimonios platónicos, esporádicos pero muy significativos, den a entender que existe una asimetría estructural irreductible entre la experiencia del bien y la del mal: mientras que el bien se experimenta como tal, sobre todo, en el modo de la identificación que caracteriza a la perspectiva de la primera persona, allí donde se escoge lo que se considera bueno, la experiencia del mal *como mal* supone, en cambio, el distanciamiento respecto de aquello que es considerado como malo, de modo tal que considerar y experimentar el mal *como mal* significa, más bien, no identificarse con él y, por tanto, no

¹¹ El mismo diagnóstico acerca de la inevitabilidad de la referencia a un componente básico de error cognitivo en la explicación de la opción por el mal parece subyacer a la posición asumida por Terence Penner, en su defensa de la explicación reductiva de la incontinencia que el Sócrates platónico presenta en el *Protágoras*. Penner intenta mostrar cómo la posición de Sócrates, adecuadamente reconstruida, explica el fenómeno de la incontinencia de modo más convincente que modelos explicativos basados en la división psicológica, en la línea del Platón maduro y de Davidson. Véase Penner, 1990.

escogerlo, esto equivale a decir que debe considerarse siempre exclusivamente desde la perspectiva externa que corresponde a la segunda o la tercera persona.¹² En esta misma dirección parece apuntar también la sugestiva recomendación de San Pablo, cuando señala a los cristianos de Corinto que se hagan expertos en las cosas del espíritu, que requieren del buen juicio, y permanezcan, en cambio, inexpertos como los niños en las cosas del mal (*cfr. Corintios I, 14, 20*).

En cualquier caso, y sea como fuere que se quiera responder al dilema planteado por la posición socrática relativa al error moral, un segundo aspecto en que la influencia de la aproximación practicada por Sócrates

¹² En un sugestivo pasaje del *Protágoras*, comparando la función que los alimentos desempeñan respecto del cuerpo con la que los conocimientos o, en general, los contenidos de creencia desempeñan respecto del alma, Sócrates enfatiza la mayor peligrosidad de estos últimos: quien compra alimentos puede no ingerirlos directamente, sino llevarlos en un envase, de modo que no lo dañen, si llegan a estar en mal estado o ser perjudiciales para la salud; en el caso de los contenidos de creencia, en cambio, tal tipo de mediación ya no resulta posible, pues el alma los incorpora de modo inmediato y los asimila sin más. Esto quiere decir que en todo acto de apropiación de un contenido de creencia tiene lugar, de manera inevitable, una cierta identificación con dicho contenido, que impide reconocer, de allí en más, su carácter eventualmente falso y perjudicial. Por ello, Sócrates advierte al joven Hipócrates, deseoso de convertirse en discípulo del sofista Protágoras, que los recaudos que impidan una identificación inmediata con contenidos cualesquiera no sometidos a examen resultan aquí más necesarios todavía que en el caso de los alimentos (*cfr., Protágoras, 314a-c*). En un sentido comparable, en la segunda prueba de la superioridad del filósofo frente al hombre tiránico sometido a los placeres sensibles de *República IX*, Sócrates explica que hay tres clases de placeres, correspondientes a las tres partes del alma y a los tres tipos de hombre que viven según cada una de ellas, a saber: los placeres del intelecto propios del filósofo, los placeres del honor y el poder propios del hombre de coraje, y los placeres sensibles propios del hombre que vive según sus apetitos sensibles. Ahora bien, Sócrates sostiene, aunque cada uno de ellos los frecuenta en diferente medida y de diferente modo, hay un único hombre que conoce mejor los tres tipos de placeres: el propio filósofo, en la medida en que es él quien está en condiciones de ponderar adecuadamente la posición relativa que merece cada uno de ellos, en una vida ordenada con arreglo a los patrones de racionalidad que reflejan el ideal de la justicia (*cfr., República, 580d-583a*). Esto implica, una vez más, que el genuino conocimiento de lo

resulta decisiva y perdurable concierne al reconocimiento expreso de la indispensable función positiva que desempeñan los factores de contraste y alteridad, a la hora de dar cuenta de la posibilidad de superar el autoengaño que caracteriza a las diferentes modalidades de la conciencia errónea, a través de las correspondientes experiencias de autodistanciamiento. En este punto, como se vio, las posiciones en la línea de Platón, Aristóteles y Tomás de Aquino, que apuntan a explicar el conflicto motivacional por recurso a diferentes modelos de división psicológica, pueden verse incluso como el resultado de un esfuerzo consciente por conservar y radicalizar el punto de vista alcanzado por Sócrates, proyectándolo, desde la esfera exterior del diálogo y la interacción social, también a la esfera interior del alma individual, concebida ahora como una entidad compleja que alberga en sí una multiplicidad de fuerzas motivacionales, capaces de entrar en conflicto unas con otras. En atención al papel positivo que cumplen los factores de contraste y alteridad, como medios que hacen posible la superación del autoengaño, puede decirse, siguiendo la lúcida conclusión de Davidson, que la explicación de la posibilidad de la irracionalidad interna y la de la posibilidad del progreso moral deben verse, en definitiva, como el anverso y el reverso de un mismo problema básico, en el marco de una teoría que apunte a dar cuenta de la estructura de nuestras acciones, como agentes racionales.¹³

No ha de resultar demasiado sorprendente, dado el carácter inigualablemente paradójico de su figura y su pensamiento, que haya sido precisamente Sócrates, el filósofo más reacio a aceptar la posibilidad de la irracionalidad interna, quien más nos haya ayudado a comprender la imprescindible función positiva que desempeñan, en el caso de un ser finito y falible como el hombre, los fenómenos de conflicto, de división interior y de toma de distancia crítica respecto de sí mismo.

que ha de verse como malo o perjudicial presupone en esencia un componente de distanciamiento, como el que caracteriza a la perspectiva externa de la segunda o la tercera persona.

¹³ Véase Davidson, 1982: 305; 2004a: 187. "A theory that could not explain irrationality would be one that also could not explain our salutary efforts, and occasional successes, at self-criticism and self-improvement".

BIBLIOGRAFÍA

- Aristóteles (1894), *Ethica Nicomachea* (EN), editado por I. Bywater, Oxford, Estados Unidos, Oxford University Press.
- Davidson, Donald (2004a), "Paradoxes of irrationality", en Donald Davidson, *Problems of Rationality*, Nueva York, Oxford University Press, pp. 169-187.
- Davidson, Donald (2004b), "Incoherence and irrationality", en Donald Davidson, *Problems of Rationality*, Nueva York, Oxford University Press, pp. 189-198.
- Davidson, Donald (1985), "Incoherence and irrationality", *Dialéctica*, vol. 39, núm. 4, pp. 345-354.
- Davidson, Donald (1982), "Paradoxes of irrationality", en Richard Wollheim y James Hopkins (eds.), *Philosophical Essays on Freud*, Cambridge, Massachusetts, Cambridge University Press, pp. 289-305.
- Davidson, Donald (1980 [c. 1970]), "How is weakness of the will possible?", en Donald Davidson, *Essays on Actions and Events*, Nueva York, Oxford University Press, pp. 21-42.
- Gómez Lobo, Alfonso (1998), *La ética de Sócrates*, Santiago de Chile, Andrés Bello.
- Irwing, Terence H. (ed.) (1995), *Classical Philosophy. Collected Papers*, vol. 3: *Plato's Ethics*, Nueva York/Londres, Garland.
- Mele, Alfred R. (1992 [c. 1987]), *Irrationality. An Essay on Akrasia, Self-Deception, and Self-Control*, Nueva York, Oxford University Press.
- Penner, Terence (1990), "Plato and Davidson: Parts of the soul and weakness of the will", *Canadian Journal of Philosophy*, vol. 16, suplemento, pp. 35-74. [Reproducido en Terence H. Irwin (ed.) (1995), *Classical Philosophy. Collected Papers*, vol. 3: *Plato's Ethics*, Nueva York/Londres, Garland, pp. 143-182.]
- Platón (2003), *Platonis Rempublicam*, edición de S. R. Slings, Oxford, Estados Unidos, Oxford University Press.
- Platón (1903), *Gorgias*, en *Platonis Opera*, vol. 3: *Tetralogiam V-VII Continents*, edición de John Burnet, Oxford, Estados Unidos, Oxford University Press.
- Platón (1903), *Protagoras*, en *Platonis Opera*, vol. 3: *Tetralogiam V-VII Continents*, edición de John Burnet, Oxford, Estados Unidos, Oxford University Press.
- Price, A. W. (1995), *Mental Conflict*, Londres/Nueva York, Routledge.
- Spitzley, Thomas (1992), *Handeln wider besseres Wissen. Eine Diskussion klassischer Positionen*, Berlín/Nueva York, Gruyter.
- Tomás de Aquino (1988 [c. 1962]), *Summa Theologiae*, Roma, Editiones Paulinae.
- Vigo, Alejandro G. (2002), "Autodistanciamiento y progreso moral. Reflexiones a partir de un motivo de la ética socrática", *Diadoché*, vol. 5, núms. 1-2, pp. 65-101.

- Vigo, Alejandro G. (2001), "Platón, en torno a las condiciones y la función del diálogo cooperativo", *Tópicos*, núms. 8-9, pp. 5-41.
- Vigo, Alejandro G. (2000), "Platón, en torno a las condiciones y la función del diálogo cooperativo", *Anales de la Academia Nacional de Ciencias de Buenos Aires*, vol. XXXIV, núm. 2, pp. 509-542.
- Vigo, Alejandro G. (1998), "Incontinencia e irracionalidad según Davidson", *Cuadernos del Sur-Filosofía*, núm. 28, enero, pp. 51-97.

Alejandro G. Vigo: Profesor en Filosofía (1984) y Licenciado en Filosofía (1988), por la Universidad de Buenos Aires (Argentina); Doctor en Filosofía (1994), por la Universidad de Heidelberg (Alemania). Desde 2006 es Profesor Ordinario del Departamento de Filosofía de la Universidad de Navarra (Pamplona, España). Ha sido becario del CONICET (Argentina), del Deutscher Akademischer Austauschdienst (Alemania) y la Fundación Alexander von Humboldt (Alemania). Desde 2006 es Miembro Titular del Institut International de Philosophie, École Normal Supérieure (París, Francia). En octubre de 2010 recibió el Premio Friedrich Wilhelm Bessel, concedido por la Fundación Alexander von Humboldt y el Ministerio de Educación e Investigación de Alemania, como reconocimiento a la trayectoria en la investigación. Actualmente están en prensa sus libros *Juicio, experiencia, verdad. De la lógica de la validez a la fenomenología* (Pamplona, 2013) y *Praxis, Reason and Truth. Studies in Aristotle's Conception of Praxis and Practical Rationality* (Leuven, 2013).

D. R. © Alejandro G. Vigo, México D.F., enero-junio, 2013.