

LA SINRAZÓN DE LA RAZÓN COMO ESTRATEGIA. RAZONES EN FAVOR DE LA RACIONALIDAD DE LOS AGENTES MORALES*

THE UNREASON OF REASON AS A STRATEGY.
REASONS FOR THE RATIONALITY OF MORAL AGENTS

MARÍA EUGENIA RODRÍGUEZ PALOP
Universidad Carlos III de Madrid

Fecha de recepción: 9-3-12
Fecha de aceptación: 22-3-12

Resumen: *La especificación de la razón como un fenómeno subjetivo no puede agotar, aunque lo pretenda, la definición de la razón, y ello, entre otras cosas, porque se apoya en el presupuesto irracionalista del egoísmo. La adopción de un punto de vista racional exige la incorporación de un punto de vista imparcial y la imparcialidad se alcanza sobre todo en procesos dialógicos de razonamiento. Dado que el aspecto normativo no puede eliminarse en el discurso racional, tampoco puede eludirse la conexión que existe entre racionalidad y moralidad.*

Abstract: *The specification of reason as a subjective phenomenon cannot exhaust, but it seeks, the definition of reason, and this is due, among other things, because it is based on irrational selfishness. The adoption of a rational point of view*

* La publicación de este artículo ha recibido el apoyo del Proyecto Consolider-Ingenio 2010: *El Tiempo de los derechos* (CSD 2008-00007) y del Proyecto: *Los Derechos Humanos en el S. XXI. Retos y desafíos del Estado de derecho global* (DER 2011-25114), ambos liderados por el Instituto de derechos humanos "Bartolomé de las Casas" de la Universidad Carlos III de Madrid. Se ha beneficiado también, en sus diferentes versiones, de los comentarios de Víctor Bermúdez, Juan Antonio Negrete, Rafael de Asís, Francisco Javier Ansuátegui, Marisa Iglesias, Ricardo García Manrique, Víctor Méndez, Joaquín Valdivieso, Carlo Ferrajoli y Pere Vilanova. Y es, además, deudor del debate que suscitó en el Seminario de Filosofía del Derecho de la Universitat Pompeu Fabra (Barcelona), hace ya algunos años, así como de las sugerencias de los/as alumnos/as y profesores/as que participan asiduamente del Seminario-Taller de Teoría Crítica, coordinado por Carlos Lema y Silvina Ribotta en el Instituto de derechos humanos "Bartolomé de las Casas".

requires the incorporation of an impartial standpoint and fairness is achieved especially in dialogical processes of reasoning. Since the normative can't be eliminated in a rational discourse, it cannot escape a connection between rationality and morality.

Palabras claves: racionalidad instrumental, egoísmo, neutralidad temporal, racionalidad comunicativa

Keywords: instrumental rationality, selfishness, temporal neutrality, communicative rationality

“La pregunta “¿Por qué debo guiarme por la razón?” simplemente no tiene sentido. Preguntarlo muestra una falta completa de comprensión del significado de una pregunta por el porqué [...] dado que “¿Debo guiarme por la razón?” significa “Dime si hacer lo que se fundamenta en las mejores razones es hacer lo que se basa en las mejores razones”, simplemente no hay posibilidad de agregar “¿Por qué?”. La cuestión se reduce a “Dame las razones de por qué hacer lo que se basa en las mejores razones es hacer lo que se basa en las mejores razones. Precisamente es como preguntar “¿Por qué un círculo es un círculo?””¹.

1. INTRODUCCIÓN

A pesar de las palabras de Baier expresadas más arriba, a lo largo de nuestra vida no parece que podamos dejar de preguntarnos acerca de la razón ni tampoco acerca del porqué de la razón. Y ello aunque la simple formulación de esa pregunta nos sitúe ya en un proceso racional y pudiera parecer inútil o absurda. Hay que considerar que cualquier idea que pretendiese justificar la irracionalidad o, sencillamente, animarnos a aceptarla como un devenir necesario, sería desde ese momento autocontradictoria, y no podría sugerirse sino desde la razón, o, dicho de otro modo, un cuestionamiento formulado contra el razonamiento tendrá que contener su propio razonamiento y tal razonamiento sólo podrá ser evaluado racionalmente. Pocas cosas hay tan complejas como relevantes, y a pocas se les han dado tantas respuestas insatisfactorias. Parece que seguimos condenados a preguntarnos una y otra vez qué es una exigencia de la razón o si es posible llegar a saber

¹ K. BAIER en *The Moral Point of View: A Rational Basis of Ethics*, Ithaca, Cornell University Press, New York, 1958, pp. 160-161.

qué es lo que la razón exige. Seguramente los esfuerzos que se han presentado hasta ahora no han sido por completo inútiles, pero para encontrar cualquier camino tenemos que analizar un buen número de propuestas, y algunas de ellas consisten en renunciar a la certeza como objetivo. Lamentablemente, en este artículo no voy a presentar ninguna visión completa de tales propuestas, ni pretendo, siquiera de manera superficial, abordar algo semejante. Lo que quiero hacer es sólo lo siguiente:

(I) Plantear una aproximación a la idea de racionalidad instrumental (también llamada estratégica o de medios) a fin de mostrar sus insuficiencias como concepto total de racionalidad. La especificación de la razón como un fenómeno subjetivo (en sentido particular o psicológico) mediante el que se logra poner en marcha un cálculo matemático de minimización de costes y maximización de beneficios, o la selección de los medios más eficaces para lograr ciertos fines, no puede agotar la definición de la racionalidad que, en mi opinión, no encaja plenamente en el simple ámbito de la instrumentalidad².

Debo subrayar que esta idea sólo me interesa por lo que hace a la mayor o menor virtualidad de la racionalidad instrumental en orden a la construcción de una teoría moral, por lo que me detendré únicamente en aquellas posiciones desde las que se ha defendido expresamente la vinculación entre un determinado modelo racional y un pretendido discurso moral, distinguiendo las que asumen la perspectiva de un sujeto egoísta que sólo incorpora fines subjetivos a su razonamiento, y aquellas que consiguen dar el salto desde un esquema prudencial a un punto de vista imparcial. Cuando digo esto, excluyo del análisis tanto las posiciones que no colocan el autointerés en el punto de partida, como aquellas que no se plantean la necesidad de adoptar un punto de vista imparcial.

² En este terreno, sería interesante tener presente la aportación de WEBER que, aunque opta por la racionalidad formal (instrumental) frente a la sustantiva (orientada a fines éticos), asume que la primera, desgajada de la segunda, posee un potencial sumamente dañino. De hecho, el propio WEBER parece poner en cuestión la distinción conceptual pura entre la esfera formal y la material, si bien admite la utilidad práctica y la necesidad de tal distinción y establece algunas matizaciones de interés (en *La acción social: ensayos metodológicos* (1922/1969), trad. M. Faber-Kaiser y S. Giner, Península, Barcelona, 1984, pp. 68-69 y pp. 72-76, y en *Economía y sociedad: esbozo de sociología comprensiva* (1922), Fondo de Cultura Económica, México, 1993, pp. 20-21). La forma en la que entiende la dicotomía medio/fin se refleja bien en la relación que establece entre la "ética de la convicción" y la "ética de la responsabilidad" (en *La ciencia como profesión/La política como profesión* (1919), trad. J. Abellán, Colección Austral, Madrid, 1992, pp. 149-160).

(II) Demostrar que la de un sujeto maximizador de intereses egoístas es una presunción irracionalista en la concepción misma de la racionalidad estratégica, dado que el egoísmo es inconsistente (no puede proporcionar una orientación consistente para la elección del egoísta o no puede hacerlo en todos los casos) y/o es contraproducente (no sirve para la realización de los intereses mejores del egoísta). De hecho, paradójicamente, el uso instrumental de la razón parece indicar que lo más racional es optar por un principio no egoísta de decisión.

(III y IV) Señalar que si el egoísmo se compadece mal con la racionalidad, no puede ser la base de ninguna teoría moral aceptable. Aunque pudiera hablarse de razones de autointerés para aceptar restricciones a la satisfacción del propio interés, esto, por sí solo, no permite al egoísta adoptar un punto de vista moral. De hecho, no parece plausible pasar de la racionalidad estrictamente estratégica a la teoría moral si uno no incorpora un universo de fines establecidos de manera imparcial. Voy a presentar este paso como un paso necesario y, sin embargo, insuficiente, para la elaboración de una teoría moral aceptable.

Lo que intento mostrar es que si hablamos de fines subjetivos que todos podemos asumir a lo largo de nuestra vida y nos situamos en una forma de universalismo débil (tal como ha hecho Gauthier), nos encontraremos con ciertos problemas de inconsistencia, derivados de combinar el egoísmo con la idea de la neutralidad temporal (todos los momentos de la vida están en pie de igualdad). Si hablamos de fines establecidos desde un punto de vista imparcial, si asumimos que no es posible eliminar el aspecto normativo en el discurso racional, avanzamos algo más, pero si seguimos utilizando un razonamiento monológico en la construcción de una teoría moral, el alcance de esta teoría queda notablemente reducido. Y es que, a mi juicio, la imparcialidad sólo puede alcanzarse de forma satisfactoria acudiendo a ciertos modelos dialógicos, como el del constructivismo habermasiano, por ejemplo, una vez subrayados sus elementos fácticos. De hecho, si esta perspectiva puede distinguirse de la postura del observador imparcial es únicamente por la relevancia que en ella tienen tales elementos. La cuestión, que no examinaré aquí, es saber si Habermas ha logrado corregir suficientemente su modelo como para haber pasado de un concepto clásico de imparcialidad a uno más complejo y completo; uno que le permita apostar por una teoría moral en la que pueda apreciarse, en línea con Seyla Benhabib, al “otro concreto”³.

³ S. BENHABIB, *El ser y el otro en la ética contemporánea: feminismo, comunitarismo y post-modernismo* (1992), trad. G. Zadunaisky, Gedisa, Barcelona, 2006, pp. 171 y ss.

2. LA RACIONALIDAD ESTRATÉGICA O LA REDUCCIÓN DE LA RAZÓN AL CÁLCULO ESTRATÉGICO

Puede decirse que la racionalidad instrumental exige pasar del pensamiento paramétrico, en el que el actor racional trata a su medio como una constante, a uno estrictamente estratégico, en el que el medio está integrado por otros actores cuyas intenciones se han de tomar en cuenta, incluyendo el hecho de que tales intenciones se basan en expectativas concernientes a las tuyas propias. De este modo, un ser racional (estratégico) no sólo toma sus decisiones a partir de sus expectativas de futuro sino sobre la base de sus expectativas acerca de las expectativas de los demás. En este contexto, hay que tener en cuenta que en una comunidad de actores estratégicos el destino no depende de fuerzas suprainstanciales, mientras que los actores paramétricos están sometidos a fuerzas causales que pueden ignorarlos y frustrar perpetuamente sus planes⁴. La racionalidad estratégica es la que, en tales circunstancias, lleva al sujeto a hacer un cálculo que le permita obtener el máximo beneficio para él en su interacción con otros sujetos racionales.

De este modo, la racionalidad instrumental se presenta como una racionalidad de medios y se identifica, sobre todo, con las mejores razones para las creencias, con el elemento cognoscitivo de la racionalidad, subvalorándose así tanto el aspecto práctico de la racionalidad, cuya tarea es implementar las mejores razones para las acciones (directivas), como el evaluativo (mejores razones para las evaluaciones o valoraciones). Esta visión de la racionalidad se apoya en la identificación de la razón con su vertiente lógico-matemática, marginándose los aspectos materiales o sustantivos del razonamiento. Por supuesto, quedaría pendiente explicitar el criterio de racionalidad que justifica estas distinciones o la razón última por la que los elementos cognoscitivo, práctico y evaluativo, no son entendidos como contextos interdependientes para el desenvolvimiento de la racionalidad y no como modos o tipos de racionalidad.

Parecería más acertado adivinar que la racionalidad práctica requiere de la cognoscitiva, del reconocimiento de nuestras razones, y de la evaluativa, de la evaluación de tales razones como las mejores. La racionalidad cognoscitiva necesita de la práctica, pues la experiencia ha de validar nuestras hipótesis sobre el mundo, fruto de nuestras acciones, y, también de la evalua-

⁴ J. ELSTER, *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad* (1979), trad. J.J. Utrilla, Fondo de Cultura Económica, México, 1997, pp. 38-40.

tiva, porque no es posible proceder en asuntos fácticos si no se consideran los juicios de valor. Sin tales juicios, no podríamos ni valorar nuestros fines ni escoger los medios tendentes a su consecución. Y finalmente, la racionalidad evaluativa necesita información acerca de hechos y acciones, aunque nos exija pasar de evidencias subjetivas a conclusiones que no lo sean, en sentido estricto.

Lo cierto es que desde una perspectiva reduccionista, la racionalidad instrumental tiene sobre todo un carácter descriptivo y sólo secundariamente normativo o prescriptivo. Describe cómo actuamos e indica cómo no debemos actuar, si bien sólo de forma incompleta, dado que ha de considerar la existencia de contextos de incertidumbre o ignorancia radical, así como de variables que no siempre se dan adecuadamente (la magnitud del problema, el número de agentes involucrados y las restricciones a las que se ven sometidos)⁵. De hecho, en términos normativos podría decirse que esta teoría de la racionalidad es indeterminada, impropia e inconclusa, porque recomienda la elección de los mejores medios para la consecución de los fines del agente, pero no ofrece los instrumentos necesarios para hacer el cálculo que tal elección requiere. Esto no significa, por supuesto, que sea completamente inútil, ni que haya otros modelos de racionalidad de los que, necesariamente y *a priori*, pueda predicarse lo contrario. Lo único que demuestra es que no es una teoría normativamente satisfactoria, aunque pueda serlo descriptivamente. Ciertamente, para algunos, esto no es en absoluto problemático, pero para nosotros sí que lo es. Si un concepto de racionalidad no nos indica adecuadamente qué debemos y no debemos hacer para obtener nuestros fines, no puede ser un buen criterio de racionalidad y, por supuesto, no puede servir a ninguna teoría moral aceptable. Ya lo veremos.

En definitiva, y derivado de lo anterior, tenemos que la racionalidad instrumental se mide en función de la mayor o menor capacidad del agente para escoger los medios adecuados a la consecución de sus fines (cualesquiera que estos sean, dado que ninguno puede ser considerado irracional y todos pueden ofrecer razones para actuar), por lo que parece lógico pensar que tal modelo racional se apoya, al menos, en dos datos: la probabilidad que tiene una acción de generar consecuencias y la utilidad esperada de ca-

⁵ Pueden verse a este respecto, los estudios de J. ELSTER en *Ulises y las sirenas. Estudios sobre racionalidad e irracionalidad* (1979), *cit.*, pp 15 y ss. y pp. 38-40, o en *Juicios salomónicos. Las limitaciones de la racionalidad como principio de decisión* (1989), trad. C. Gardini, Gedisa, Barcelona, 1991, pp. 16-24 y pp. 32-33, entre otros.

da una de tales consecuencias. La utilidad esperada está relacionada con los fines del sujeto (aunque no se considera aquí como un *prius* lógico o un valor constitutivo de los mismos) y, a su vez, en el esquema estratégico, tales fines están relacionados con la satisfacción de sus deseos y preferencias subjetivas⁶.

Esto significa que un sujeto racional es el que consigue orientar sus deseos a los fines que se propone (automanipulación) o bien adaptar tales fines a los deseos que tiene (determinismo psicológico); un sujeto racional es el que, en todo caso, logra hacer coincidir deseos y utilidad esperada. Dicho de otro modo, el sujeto racional es el que no desea lo imposible, ni siente frustración respecto de lo inevitable, o bien porque se adapta al medio por un proceso de planificación consciente (controla la situación), o bien porque renuncia a lo que nunca podrá tener. Y es precisamente en esta capacidad de adaptación/renuncia en la que se demuestra su racionalidad.

Parece, pues, que este modelo de racionalidad se debate entre la automanipulación, con la que renunciamos a satisfacer algunos de nuestros deseos para priorizar otros que valoramos más, y el determinismo psicológico, que nos anima a asumir nuestra primigenia condición. Y la cuestión es dilucidar si ambas posibilidades están realmente al alcance del maximizador egoísta, que es el presupuesto antropológico de este modelo de racionalidad. Dilucidar, en fin, si un sujeto egoísta puede imponer condiciones sustantivas sobre algunos de sus deseos y preferencias a fin de alcanzar ciertos fines; si podría limitar algunos de sus deseos para satisfacer otros; si tendría herramientas para jerarquizarlos, y en función de qué criterio los jerarquizaría. Y, en su caso, dilucidar también si un concepto de racionalidad puede construirse sobre la asunción del determinismo como presupuesto material⁷.

Más adelante diremos algo sobre todo esto, pero antes, preguntémosnos acerca de la racionalidad o irracionalidad del egoísmo, así como por las posibilidades que tiene el egoísta de saltar, inmune, al terreno de la moralidad.

⁶ O sea, la racionalidad se percibe, a lo sumo, como una instancia motivadora de las intenciones del agente pero en ningún caso como una condición suficiente en su determinación.

⁷ No hay espacio aquí para entrar a valorar el grado en que esta forma de entender la racionalidad podría explicar y/o solventar los supuestos de racionalidad imperfecta, especialmente, el caso que representa la debilidad de la voluntad, ni tampoco hay espacio para analizar el grado en que podría llegar a provocarlos.

3. EL MAXIMIZADOR EGOÍSTA COMO SUJETO “IRRACIONAL” DE LA RACIONALIDAD ESTRATÉGICA

Comenzaremos por subrayar que lo distintivo del egoísta no es el hecho de que actúe con el objetivo de alcanzar el mayor éxito posible, sino el de buscarlo exclusivamente para sí mismo⁸. El egoísta se dirige a su propio bienestar de manera en que los componentes del bienestar del otro quedan excluidos, y se distingue así del sujeto simplemente autointeresado, que puede convertir el bienestar del otro en su meta personal. Evidentemente, en ninguno de estos dos casos se prescinde de lo que Tugendhat, llama “egocentrismo”⁹, pero sólo en el primero, tal egocentrismo puede dar lugar a una insuficiencia racional.

Pues bien, lo que intento mostrar, en la línea de muchos otros autores, es que en situaciones estratégicas, en las que un agente concreto y consciente elige en un medio sensible a él como decisor, es decir, donde se incluyen otros agentes a los que afectan sus decisiones, el egoísmo es inconsistente (no puede proporcionar una orientación consistente para la elección o no puede hacerlo en todos los casos) y/o es contraproducente (no sirve para la realización de los intereses mejores del egoísta).

3.1. Del egoísmo a la cooperación condicionada por “razones” egoístas

Como es bien sabido, el egoísmo puede ser contraproducente (no prudencial) si se demuestra no sólo que al tratar de maximizar su valor (benefi-

⁸ “Lo que el egoísmo sostiene [...] es que la felicidad de *cada* hombre es lo único bueno, que un gran número de cosas diferentes es *cada una* la única cosa buena [...]”, lo cual, según MOORE, es una contradicción absoluta. “No podría desearse una más completa y exhaustiva refutación de ninguna teoría”, afirma el autor, en *Principia Ethica* (1903), trad. T. Baldwin, Universidad Nacional Autónoma de México, México, 1997, pp. 179-180.

⁹ Es curioso que este autor vincule el egocentrismo por el que alguien convierte su bienestar en fin último, con el reconocimiento, en el sentido de aprecio ajeno. Uno no puede darse importancia a sí mismo, ni a sus fines, dice, si no es importante para otros; no puede adquirir el sentido de su propio valor si no es valioso para los demás (si no es considerado por otros como un fin en sí mismo). Y ello porque “[...] convertimos el hecho de ser reconocidos en la sensación de ser dignos de reconocimiento” (E. TUGENDHAT, *Egocentricidad y mística. Un estudio antropológico* (1997), trad. M. Suárez Crothers, Gedisa, Barcelona, 2004, p. 51). ¿Significa esto que, paradójicamente, el egoísta depende para serlo del reconocimiento y la sincera valoración de los otros, de la opción no egoísta de los demás? Me parece que indagar en este punto, en un tono psicologizante, sería altamente provechoso.

cio) los egoístas no logran que les vaya mejor colectivamente (perjudicial para todos) sino que no logran que les vaya mejor a ellos mismos. Y esto es lo que parece poner de manifiesto el conocido Dilema del Prisionero: la decisión del egoísta puede no mostrar su auténtico interés pues existe un beneficio neto que no se produce precisamente por el hecho de que el egoísta intenta maximizar su valor. Y ello porque en el Dilema del Prisionero la solución óptima para los agentes es cooperar; la mejor para cada uno es traicionar al otro; la peor (individualmente) es ser el ingenuo que coopera mientras que el otro no lo hace; la peor situación para todos es que ninguno coopere.

Es cierto, y ya se ha señalado en muchas ocasiones, que en estos casos se adopta una perspectiva colectiva, pues si bien a todos estos sujetos egoístas les va peor de lo que podría irles, esto no significa que a cada uno le vaya peor de lo necesario. O sea, que lo único que puede deducirse de este Dilema es que cuando no hay modo alguno de influir en la decisión del otro, haga lo que haga el otro, lo mejor para cada uno es no cooperar, y esto es contraproducente desde el punto de vista colectivo, pero lo que es contraproducente desde el punto de vista colectivo, no lo es necesariamente desde el punto de vista individual¹⁰. Es esto último lo que se ha planteado a veces como una de las fortalezas del egoísta: dado que en el Dilema la motivación para la acción sigue siendo la de obtener el máximo beneficio al menor coste, el egoísmo puede ser contraproducente en según qué interacciones, pero esto no significa que sea inconsistente. Sin embargo, como más adelante intentaré mostrar, la teoría fuerte del interés propio, la que mantiene el sujeto egoísta, puede desvelar graves inconsistencias. Veamos ahora si, en efecto, el egoísmo es contraproducente sólo en algunas interacciones o lo es, más bien, en todas, tal como aquí voy a sostener.

Lo que muestra el Dilema del Prisionero es que la existencia de un conjunto de acciones en equilibrio es una condición necesaria para que una decisión egoísta bien informada tenga éxito, por lo que el principio de decisión egoísta resulta insuficiente en las interacciones estratégicas en la que no existe un conjunto mejor de respuestas mutuas (un equilibrio de Nash). En

¹⁰ La dificultad de trasladar el criterio de maximización individual a acciones colectivas y la falacia en la que puede incurrirse, fueron ya demostradas por M. OLSON en *La lógica de la acción colectiva. Bienes públicos y la teoría de grupos* (1965), Limusa, México, 1992, pp. 1-2. Al respecto, puede consultarse también J. ELSTER, *Tuercas y tornillos. Una introducción a los conceptos básicos de las ciencias sociales* (1989), trad. A. Bonanno, Gedisa, Barcelona, 1990, pp. 127-135.

estos casos, los egoístas sólo pueden elegir mejores respuestas esperadas si algunos se forman expectativas erróneas¹¹.

Con todo, se dice, si se jugase el juego del prisionero a largo plazo y con el mismo compañero, un sujeto egoísta podría llegar a cooperar: si uno siempre coopera, aunque al principio fracase, el otro termina confiando en su actitud y coopera también. O sea, que a largo plazo (cuando se introduce el factor tiempo) la cooperación gana plausibilidad, y el egoísmo deja de ser contraproducente. Como dice Gauthier, el principio de decisión egoísta funciona siempre que se sustituya la selección de acciones por la de estrategias (distribución de probabilidades entre acciones posibles¹²). De hecho, en estos casos, en los que se da un número finito de agentes y estrategias, Nash ya demostró que siempre hay, al menos, un conjunto de estrategias en equilibrio, por lo que la decisión egoísta puede ser maximizadora.

Evidentemente, para que esto funcione es necesario un cierto esfuerzo de coordinación, puesto que si tal coordinación no se diera, el valor esperado del desertor tendría que ser menor. Y es aquí donde el egoísta tropieza con los primeros problemas. ¿No será tal coordinación demasiado costosa para él? Consideremos que esta coordinación pasa, como mínimo, por identificar las disposiciones del otro (algo que no es posible si no existen principios compartidos que nos permitan comprenderlas) y por dar a conocer las suyas propias, para lo que se precisa desarrollar una cierta pericia.

Salvados estos primeros escollos, lo cierto es que en estas condiciones, es muy probable que el cálculo de las consecuencias de la acción egoísta se traduzca en un derroche de tiempo y esfuerzo. Es más, incluso cuando todos los hechos

¹¹ J.F. NASH, "The Bargaining Problem", *Econometrica*, núm. 18, 1950, pp. 155-162. En los juegos estratégicos los jugadores tienen que realizar decisiones simultáneas, de tal modo que el resultado de cada una depende de las decisiones de todos. Aunque en un principio se pensó que esta interdependencia llevaría a una regresión al infinito, la noción de punto de equilibrio vino a solucionar el problema. Se define como un conjunto de elecciones tales que ningún jugador puede mejorar su resultado desviándose de él unilateralmente. El punto de equilibrio puede consistir en una estrategia pura (cada jugador debe ejecutar una determinada acción) o mixta (sorteo equilibrado de estrategias puras). También los juegos no cooperativos, en los que cada jugador decide independientemente del otro, tienen un punto de equilibrio. En los juegos de suma cero de dos jugadores, la ganancia de uno es siempre la pérdida del otro. Aquí las estrategias de equilibrio son las estrategias *maximin* de los jugadores, de modo que la cauta opción de retroceder es también estratégicamente racional.

¹² El autor introduce aquí el concepto de decisión dinámica (decisión sobre cursos de acción a lo largo del tiempo) que se refiere a intenciones sobre acciones futuras.

fueran conocidos, como señala Mackie, habría serios problemas para ligar tales hechos y sus consecuencias, sobre todo, cuando cada acto se solapa con otros muchos, existen factores condicionantes fuera del control del agente, y algunas de tales consecuencias son el producto de las elecciones de otros, independientes o no, simultáneas o posteriores, a las propias. Hablar en estos supuestos de “consecuencias probables” no aporta mayor claridad dado que habría que “sopesar una larga y completamente oscura serie de “consecuencias probables””¹³.

Por otra parte, como las acciones propositivas del egoísta serán analizadas en un entorno social (una interacción estratégica), estas acciones han de presuponer “algún nivel de regularidad y confianza en la conducta de otras personas”; es decir, que el egoísta ha de suponer que las personas responden de una manera normalizada y relativamente invariable, y que los demás harán presuposiciones idénticas a las suyas¹⁴. Por consiguiente, en la práctica, acabará por encontrar grandes ventajas en escoger principios predeterminados para regir su acción¹⁵.

Pero, ¿puede el egoísta responder satisfactoriamente a la exigencia de regularidad que lleva implícita esta forma de interacción?, ¿puede asumir principios generales y predeterminados para regir su acción? ¿No sería más racional para él abandonar este costoso camino para valorar su presente por encima del futuro que le espera? De hecho, no puede darse por descontado que el egoísta asumirá la tesis de la neutralidad temporal (todos los momentos de su vida tienen el mismo valor para él, por lo que sacrificarse hoy convirtiéndose en un cooperador podría obedecer a un cálculo egoísta de cara a la obtención de ciertos beneficios en el futuro), y el problema es que si no la asume el esfuerzo de la cooperación no le resultará rentable. Nótese que en estos casos, el egoísta ha de actuar peor para hacerlo mejor a lo largo del tiempo y si sus razones para actuar sólo proceden de sus deseos presentes reales o contrafácticos, la frustración eventual de sus predecibles deseos futuros no le dará ahora una razón para sacrificarse¹⁶. Si apelamos a estos deseos futuros, afir-

¹³ J.L. MACKIE, *Ética: la invención de lo bueno y lo malo* (1977), trad. T. Fernández Azúa, Gedisa, Barcelona, 2000, p. 176.

¹⁴ Ibidem.

¹⁵ J.L. MACKIE, *Ética: la invención de lo bueno y lo malo*, cit., p. 177.

¹⁶ Es inevitable ver aquí la estela de HUME en su *Tratado sobre la naturaleza humana* (1739-1740), Libro 2, parte 3, sección 3 (trad. F. Duque, 3 vols., Ediciones Folio, S.A., Barcelona, 2000). En esta línea es en la que se sitúan los que defienden que el *free rider* es el sujeto racional paradigmático. Sin embargo, aunque se trata de una postura interesante si lo que se pretende es mantener el egoísmo como presupuesto racional, ha sido ya suficientemente demostrada la frustración a la que está abocado el *free rider* en relaciones proyectadas en el tiempo.

mando que pueden darle al egoísta una buena razón para actuar peor ahora, estamos apelando a las razones que tiene para preocuparse e indicando, además, que debe racionalmente hacerlo, y esto es una muestra de que ya no manejamos una teoría basada en deseos o preferencias propias, que es la que da sustento al egoísmo, sino, más bien, una teoría basada en valores.

Dicho de otro modo, en una interacción estratégica proyectada en el tiempo, la cooperación no ganará plausibilidad si el egoísta no encuentra “razones egoístas” que le animen a asumir el esfuerzo que supone la coordinación con los demás a fin de alcanzar sus objetivos propios. Y el egoísta no hará tal cosa si no se compromete con la tesis de la neutralidad temporal (que es la que le permite automanipularse para obtener beneficios a largo plazo). De hecho, aún cuando el egoísta pudiera llegar a conocer (en términos ideales) todas las consecuencias probables de sus actos y tuviera una visión completa de la interacción en la que se encuentra, aún cuando pudiera llegar a racionalizar sus deseos y preferencias a fin de garantizar una cierta regularidad y pudiera confiar en que los demás harán lo mismo, al mantener un punto de vista temporalmente relativo, volvería a fracasar. Fracasaría en el terreno de la racionalidad y, por la misma razón, y como veremos después, también en el de la moralidad.

3.2. Las inconsistencias de las “razones” egoísta. La frustración inevitable

Hasta el momento, hemos visto que el objetivo maximizador del sujeto egoísta se puede alcanzar mejor mediante acciones no maximizadoras, por lo que puede ser más racional, o bien, como afirma Gauthier, ser un “egoísta incompleto” / “cooperador egoísta”¹⁷, o bien, como aquí defenderé, renunciar genuinamente al egoísmo.

Pues bien, para empezar, podemos decir con Parfit, que la teoría del propio interés incurre en algún error cuando a) nos sugiere que creamos en una teoría distinta (que nos determinemos a tener esta creencia-autorrestricción), b) que tal creencia está justificada, y c) que la teoría resultante será la mejor teoría. Y el argumento falla porque si es la mejor entonces no puede venir apoyada por la teoría del propio interés. En otras palabras, la teoría del propio interés no puede aconsejarnos cambiar de teoría acerca de la racionalidad y seguir siendo una buena teoría¹⁸.

¹⁷ Véase: *La moral por acuerdo* (1986), trad. A. Bixio, Gedisa, Barcelona, 1994. Al respecto pueden consultarse las páginas 255-306.

¹⁸ D. PARFIT, *Razones y personas* (1984), trad. M. Rodríguez González, Mínimo Tránsito/A. Machado Libros, Madrid, 2004, pp. 84-85.

El problema es, seguramente, que la teoría del propio interés se orienta a un objetivo irracional, porque puede marcar fines sustantivos en sentido subjetivo (que nuestra vida vaya lo mejor posible), y no simplemente fines formales (que seamos racionales). Veamos.

En cualquiera de sus versiones, la teoría del interés afirma:

- a) que la felicidad y el placer son, como mínimo, parte de lo que hace que nuestra vida vaya mejor;
- b) que al decidir lo que es mejor para alguien debemos dar igual importancia a todas las partes del futuro de esa persona (tesis de la neutralidad temporal)¹⁹;
- c) y que nos determinemos, o bien a tener una disposición racional con la que consigamos que nuestra vida vaya mejor (sea o no finalmente racional), o mejor, a sustituir la teoría del propio interés por otra cualquiera que nos permita alcanzar nuestros objetivos²⁰. O sea, una vez más, lo que sugiere la teoría del propio interés es que puesto que es racional para mí determinarme a ser irracional (a hacer lo que sé que es peor para mí), debo hacerlo, debo ser racionalmente irracional²¹. En otras palabras, debo actuar irracionalmente, aunque no por ello asumo que soy un ser irracional.

En consecuencia, la racionalidad está determinada por los propósitos del agente y parece que ser racional no es siempre lo que conduce a lograrlos. Luego no es siempre racional deliberar racionalmente pues no siempre conduce a un resultado racional. Incluso podría pensarse en que lo más racional es no deliberar en absoluto. Cosa distinta es que al agente le interese más el efecto general de emplear determinados procedimientos deliberativos que los resultados particulares que sus procedimientos recomiendan; es decir, que la deliberación racional sea, en sí misma, interesante, independientemente del éxito de sus resultados²². En este caso, lo importante serían las razones que tenemos para formarnos las intenciones y no sólo las accio-

¹⁹ D. PARFIT, *Razones y personas*, cit., pp. 59-62.

²⁰ D. PARFIT, *Razones y personas*, cit., pp. 66-72.

²¹ D. PARFIT, *Razones y personas*, cit., p. 75.

²² D. GAUTHIER, *Egoísmo, moralidad y sociedad liberal*, trad. P. Francés, Paidós, Barcelona, 1998, pp. 133-141. Una deliberación es racional cuando en ella se selecciona una acción que es intencionalmente compatible con las acciones previas del agente, incluyendo, entre ellas, la que forma parte del mejor curso de acción posible (instrumental) (*op.cit.*, pp. 144-159).

nes resultantes de las mismas, pero el esquema del razonamiento, en lo sustancial, seguiría siendo el mismo.

Como es fácil ver, una vez más, el problema que presenta esta teoría en orden a la posibilidad del autoengaño, tiene relación con su apuesta por la tesis de la neutralidad temporal. Ya dijimos que esta teoría da igual importancia a todas las partes del futuro de una persona por lo que no es sensible a los deseos futuros que dependen de un cambio en los juicios de valor o en los ideales propios, ni considera que puede haber buenas razones para preocuparse más por el yo en el presente que por yo en el futuro lejano (con el que tenemos una identificación menor)²³. Pero, ¿está justificada esta apuesta en una teoría fuerte del interés propio, en una teoría fundada en deseos?

Como ya hemos dicho, si estamos frente a un sujeto egoísta, no se puede descartar que lo que realice mejor sus deseos presentes no coincida con lo que de manera más efectiva promueve su interés propio a largo plazo²⁴. Y el egoísta no se sentirá motivado para optar por el futuro asumiendo su actual frustración. ¿Debería, pues, automanipularse para comprometerse con la tesis de la neutralidad temporal? Podría hablarse aquí de la existencia de *razones instrumentales dadas al estado* (como diría Pafit). Las *razones dadas al estado* no proceden del objeto de lo que queremos (*razones dadas al objeto*), sino del querer algo en sí mismo. En este caso, y tratándose, como se trata, de un sujeto egoísta, no hablaríamos de razones intrínsecas (queremos querer algo en sí) sino instrumentales, dado que lo que queremos querer lo queremos por los efectos que provoca. El egoísta quiere querer asumir la tesis de la neutralidad temporal porque esto será bueno para él.

Sin embargo, cabría preguntarse si las razones instrumentales, aún incorporadas a la larga cadena de razones instrumentales que aquí se manejan, no derivarían, finalmente, en una razón intrínseca. Evidentemente, desde la posición que defiende el egoísmo, desde una teoría fundada en deseos, al final de la cadena sólo habrá un deseo intrínseco, un deseo que no tendríamos ninguna razón para querer. O sea, que como señala Parfit, habría que concluir, finalmente, que “nuestro tener deseos no puede, por sí mismo, darnos ninguna razón”, ¿es esto posible? Si es así, sin duda, “las teorías fundadas en deseos tienen los pies de barro”²⁵.

²³ En *Razones y personas* (1984), *cit.*, pp. 537-558.

²⁴ *Idem*, pp. 784-792.

²⁵ *Persona, racionalidad y tiempo*, trad. J.O. Benito Vicente, Síntesis, Madrid, 2004, p. 74.

La cuestión es que estas teorías ignoran las razones intrínsecas que existen para querer algo o para querer querer, así como la misma idea de racionalidad intrínseca²⁶, y es por esto por lo que el egoísta sufre un problema de insuficiencia racional que no le permite conseguir lo que se propone. Lo más coherente sería que el egoísta asumiera que hay deseos (ideales, valores, creencias) que son irracionales (o están abiertos a la crítica racional) y que no proporcionan buenas razones para actuar, mientras que hay deseos que son racionalmente requeridos (hay, pues, razones que no dependen de deseos)²⁷. Dicho de otro modo, si los deseos no suelen proveer razones por sí solos sino que están apoyados en razones externas, si son las razones que tenemos para actuar de un modo concreto las que explican nuestro deseo de actuar de ese modo²⁸, no hay ningún motivo *a priori* para excluir la opción por el presente. La cuestión, insisto, es que el egoísta no puede asumir tal cosa. Si los deseos han de supeditarse a razones externas y jerarquizarse conforme a tales razones, no se ve cómo puede el egoísta recurrir a “razones egoístas” (a

²⁶ *op.cit.*, pp. 70-77. Para PARFIT, “algo es deseable si tiene características que nos dan razones para quererlo” (*op.cit.*, p. 74). Ya veremos que esta perspectiva también puede ser contestada.

²⁷ D. PARFIT en *Razones y personas* (1984), *cit.*, y en *Persona, racionalidad y tiempo*, *cit.*, pp. 74-77.

²⁸ PARFIT apuesta, así, por una teoría fundada en valores (*value-based*), según la cual las razones prácticas son proporcionadas por hechos acerca de lo que merece la pena conseguir, y rechaza expresamente las teorías fundadas en deseos (*desire-based*). Hay que notar que cuando Parfit habla de algo que merece la pena conseguir no se está refiriendo a lo bueno en sí mismo, sino a algo que tiene ciertas propiedades que lo hacen bueno (bondad instrumental), y esto ya distingue su posición de la de autores que, como Moore o Nagel, también apuestan por una teoría fundada en valores (más consecuencialista). Aunque Parfit asume que las verdades acerca de las razones son normativas, no incluye entre ellas verdades acerca de lo que es bueno o malo; es decir, no necesita apelar a la propiedad no natural de la bondad. A este respecto, puede consultarse: *Persona, racionalidad y tiempo*, *cit.*, pp. 63 y ss. Las diferencias entre Parfit y Nagel, en este y en otros campos, son, por lo demás, evidentes. Así como Parfit niega rotundamente y hasta la saciedad la existencia de egos cartesianos, Nagel se declara abiertamente cartesiano, y asume que hay ciertos pensamientos respecto de los cuales no podemos ubicarnos *en el exterior* porque están incluidos inevitable y directamente en cualquier procedimiento que nos lleve a considerarnos desde un punto de vista externo (por ejemplo, en *La última palabra. La razón ante el relativismo y el subjetivismo* (1997), trad. P. Bargallo y M. Alegre, Gedisa, Barcelona, 2000, pp. 30-36). Estas diferentes fuentes filosóficas marcan definitivamente la posición que cada uno mantiene en relación con la racionalidad y la teoría moral. En este punto, Parfit parece conectar con W. QUINN en “Putting Rationality in Its Place”, *Morality and Action*, Cambridge University Press, 1993.

sus propios deseos) para automanipularse y alcanzar así sus objetivos. O sea, que la teoría instrumental falla en el terreno de la racionalidad y obliga al egoísta a dejar de serlo, en un sentido genuino y no por razones egoístas, sino por buenas razones, por razones que gozan de una racionalidad intrínseca. Y de esta manera se muestra la inconsistencia del egoísmo, dado que no hay razones egoístas para *ser* un egoísta (no ya para actuar como un egoísta). El egoísmo es contraproducente en todo tipo de interacciones (también a largo plazo) y es, además, inconsistente.

Podría decirse, como señaló Parfit hace años, que aunque la teoría instrumental fracase en el tiempo (es intertemporalmente autoanulante) esto no dice nada en su contra porque no pretende ser intertemporal²⁹. Pero esta aclaración no afecta en absoluto a lo que aquí nos interesa mostrar: que una teoría temporalmente relativa y estrictamente instrumental no ofrece al egoísta ninguna motivación para seguir siéndolo, ni para convertirse en un cooperador (autoengañarse a fin de obtener beneficios a largo plazo), ni para transformarse en un agente moral, sino que sólo puede animarle a rechazar, por irracional, semejante teoría. Y es que, tanto si estamos frente a una interacción proyectada en el tiempo como si no es este el caso, esta teoría que anima al egoísta, le conduce irremisiblemente a un fracaso; un fracaso que demuestra tanto su insuficiencia racional como sus carencias morales.

4. LA RELEVANCIA DE LA TESIS DE LA NEUTRALIDAD TEMPORAL EN EL SALTO DEL ESQUEMA PRUDENCIAL AL TERRENO DE LA MORALIDAD

Está claro que la posición que Parfit defiende, de la que acabamos de dar cuenta muy brevemente, resulta más racional (en términos de autointerés) que la del egoísmo, dado que apuesta desde un principio por la continuidad psicológica del sujeto y por la racionalización de su experiencia³⁰. Sin tal racionalización sería imposible jerarquizar los deseos y los fines propios, y ello aun cuando haya de asumirse un cierto grado de variabilidad. Además, Parfit abre la puerta a una teoría moral más aceptable, en la que se adopta el punto de vista interper-

²⁹ D. PARFIT, *Prudencia, moralidad y dilema del prisionero* (1981), trad. G. Gutiérrez, Editorial Complutense, Madrid, 2007, p. 36.

³⁰ Para hablar de continuidad psicológica, PARFIT apela a “[...] diversas relaciones psicológicas implicadas en la memoria o en la persistencia de intenciones, deseos y otros rasgos psicológicos” (en *Persona, racionalidad y tiempo*, cit., p. 99).

sonal como fruto de la conectividad psicológica que existe entre los sujetos (cuando el yo pierde relevancia y me preocupa menos el resto de mi vida, los otros están más cerca de mí). Digamos, por contraste, que la moralidad del egoísta se presentaba como un conjunto de restricciones a la maximización de su valor-relativo, de tal manera que, dado que aspiraba a maximizar su valor, encontraba racional someterse a ellas. O sea, que la moralidad se presentaba, coherentemente, como el fruto de una supuesta (imposible) automanipulación racional, y los principios morales como principios para la decisión racional³¹.

Sin embargo, la construcción de Parfit, avanzando algunos pasos, acaba planteando también dudas.

Si los deseos tienen su base en razones externas, derivadas de nuestra experiencia (nuestro pensamiento está condicionado por los hechos que vivimos y de los que somos conscientes); o, dicho de otro modo, si nuestras creencias están causalmente determinadas, ¿es este determinismo racionalmente aceptable?

Aquí parece útil la distinción que hace Mackie entre determinismo débil y fuerte. En el primero se asume lo que este autor llama "compatibilismo", el hecho de que el determinismo puede aceptarse sin modificar por ello nuestro pensamiento moral. En el segundo, se combina el determinismo con el "incompatibilismo", con la incompatibilidad entre determinismo y nociones morales esenciales³². El determinismo débil no es fatalismo, ni identifica causación y obligación, ni nos ofrece una explicación mecánica del modo en que actuamos, por eso puede ser "compatibilista".

³¹ D. GAUTHIER, *Egoísmo, moralidad y sociedad liberal*, cit., p. 93. Si pensamos en lo que hemos visto hasta ahora, podríamos decir que aquí el término "moral" se emplea sólo en sentido descriptivo, y no se identifica con lo "moralmente bueno". O sea, que se plantea en aquellos casos en los que supuestamente podemos afirmar que algo es "moral" y no comprometernos con ello, de modo que, incluso, podríamos llegar a condenarlo. La cuestión es que lo que aquí defendemos es que los términos morales poseen simultáneamente un significado descriptivo y normativo, y es porque asumimos esto por lo que la posibilidad de universalizar es una tesis lógica. Evidentemente, hay quienes, como MACKIE, por ejemplo, no deducen de la (supuesta) verdad de esta tesis lógica la obligación de pensar así, aún a riesgo de actuar sin lógica. La verdad de una tesis lógica, dice el autor, "no obliga en modo alguno a la aceptación del principio práctico sustantivo correspondiente", dado que siempre se precisa una decisión ulterior a favor o en contra de este principio y cada uno es lógicamente libre de efectuar su opción fuera del juego moral (en *Ética: la invención de lo bueno y lo malo* (1977), cit., pp. 112-113). Lo cierto es que no consigo asumir lógicamente esta afirmación.

³² Idem, p. 251.

Los deterministas débiles entienden que la consideración y la valoración de alternativas forma parte del proceso causal, y que el carácter de una persona puede moldearse a la luz de sus experiencias. Sin embargo, si esto es así, tanto la valoración de alternativas, como la modificación del carácter como fruto del aprendizaje, en buena lógica, ha de venir también causalmente determinado, y si lo está también por las experiencias, el razonamiento podría ser circular, con lo que no ofrecería ni una explicación completa de nuestras acciones, ni, por supuesto, una justificación. ¿Socava esto la racionalidad de las elecciones y su alcance moral? Como señala Mackie, esta postura requiere “un yo contracausal libre y sin embargo determinado y activo”, y este es un concepto muy difícil de aprehender coherentemente³³.

Finalmente, el determinismo débil equipara el yo con un sistema de rasgos empíricos contingentes, por lo que esta perspectiva, por un lado, va unida a la negación de una identidad absoluta en el tiempo³⁴, y, por el otro, nos ofrece una descripción de la razón como fenómeno natural (biológico o psicológico) pero nos deja sin ninguna “buena” razón para confiar en ella. El problema es que cuando la identidad se reduce a hechos, psicológicos o de cualquier tipo, corre el riesgo de dar por supuesto lo que quiere demostrar³⁵. Como señala Nagel, “[...] no podemos incluir en una teoría psicológica todo nuestro razonamiento, incluyendo los razonamientos que nos condujeron a dicha teoría psicológica”. Obviamente, esta postura no ha de llevarnos a afirmar que hay premisas eternamente irrevisables, sino que en cualquier proceso de razonamiento debe haber algunos pensamientos que concibamos desde adentro, en lugar de considerarlos disposiciones psicológica o biológicamente determinadas. El tema es que este determinismo débil podría convertir al sujeto autointeresado no egoísta que venimos describiendo, en un sujeto imperfectamente racional.

³³ Ibidem, p. 257.

³⁴ Ibidem, p. 254. Tal es el caso de D. PARFIT, que, en no pocas ocasiones, ha venido a subrayar la irrelevancia de la identidad (por ejemplo, en *Persona, racionalidad y tiempo*, cit., pp. 93-124). En su perspectiva, lo que realmente importa son las relaciones de continuidad y conectividad psicológicas que se mantienen entre las diferentes partes de la vida de una persona. Estas relaciones pueden graduarse y coinciden sólo aproximadamente con la identidad personal (*Op.cit.*, p. 108). Es fácil ver en esta posición psicologista los ecos de J. LOCKE en su *Ensayo sobre el entendimiento humano* (1690) (Fondo de Cultura Económica, México, 1994, Cap. XXVII, libro 2º: “Acerca de la identidad y de la diversidad”) donde se hace depender la identidad de la autoconciencia. El propio Parfit, alude a Locke en algunos de los pasajes de su obra.

³⁵ En *La última palabra. La razón ante el relativismo y el subjetivismo* (1997), cit., p. 149.

No puedo entrar ahora en el debate sobre el determinismo o sobre la identidad personal y su mayor o menor relevancia en el terreno de la racionalidad (“el yo no es lo que importa”, dice Parfit), lo que me interesa es destacar que cuando se debilita la tesis de la neutralidad temporal, como hace Parfit, se acaba optando por una racionalidad imperfecta, y, por tanto, se paga un precio muy alto en el terreno de la moralidad.

Así, si se apuesta por una identidad variable y por la conformación de la identidad como el fruto de nuestras experiencias psicológicas subjetivas³⁶, pueden desvanecerse las razones para pensar en planes de vida, ya que desaparece la vinculación fuerte con nuestros yo es futuros³⁷, y, en consecuencia, también con los otros yo es que no están aquí y ahora. Como apunta Nagel, esta teoría podría tener dificultades para plantear una acción motivacional a distancia, tanto a través del tiempo como entre personas, y esto tiene una relevancia nada desdeñable por lo que hace a la adopción de un punto de vista imparcial y a la construcción de una teoría moral. De hecho, visto así, el sujeto podría incluso tener dificultades para comportarse como una unidad, dado que no podría conciliar fácilmente el punto de vista personal con el punto de vista impersonal³⁸. Y precisamente por esta razón, porque falla la

³⁶ En el paradigma kantiano se diría que la autonomía personal no se puede identificar con el solipsismo, pues exige al individuo que sea algo más que un lugar de preferencias subjetivas, y le obliga a asumir la perspectiva de la universalidad. Véase: *Crítica de la razón práctica* (1778), versión castellana de R.R. Aramayo, Alianza Editorial, Madrid, 2004, [A63-A65], pp. 106-107, y *La Metafísica de las Costumbres* (1797), trad. A. Cortina y J. Conill, Tecnos, Madrid, 1999, pp. XXVI-XXXI). No olvidemos que Kant rechaza al sujeto egoísta precisamente porque no puede alcanzar la universalidad inherente a la verdadera racionalidad. Véase: *Fundamentación de la metafísica de las costumbres* (1785), versión castellana de R.R. Aramayo, Alianza Editorial, Madrid, 2002, capítulo segundo [A88-A89], pp. 132-133 y *Crítica de la razón práctica* (1778), versión castellana de R.R. Aramayo, Alianza Editorial, Madrid, 2004, [A72-A87], pp. 114-127.

³⁷ Así lo apunta J.C. BAYÓN en *La normatividad del derecho: deber jurídico y razones para la acción*, Centro de Estudios Constitucionales, Madrid, 1991, pp. 104-105. Para una posición que defiende que todos los tiempos están en pie de igualdad, véase T. NAGEL en *La posibilidad del altruismo* (1970), Fondo de Cultura Económica, México, 2004, pp. 67-85.

³⁸ Véase T. NAGEL en *La posibilidad del altruismo* (1970), *cit.*, pp. 37-55 y en *La última palabra. La razón ante el relativismo y el subjetivismo* (1997), *cit.*, capítulo 6. Sobre el modo en que sería más o menos irrefutable el punto de vista de la primera persona, considérese aquí la conocidísima postura de B. WILLIAMS en el capítulo 4 de su *Ethics and the Limits of Philosophy* (Harvard University Press, 1985) o en su ensayo “Internal and External Reasons” publicado en su colección *Moral Luck*, Cambridge University Press, 1981.

dimensión intersubjetiva de eso que cree ser, tampoco podría convertirse en un agente moral³⁹.

Parece, pues, que la adopción de un punto de vista moral ha de empezar por defender tanto la tesis de la neutralidad temporal como la de la neutralidad interpersonal, haciendo depender la segunda de la primera. A partir del momento en que uno asume que es un yo entre otros (diferentes de lo que soy yo y a la vez iguales a mí), dice Nagel, debe ser posible decir de otras personas cualquier cosa que uno pueda decir de sí mismo y, en el mismo sentido, debe ser posible que uno se considere a sí mismo impersonalmente (aunque para lograrlo es imprescindible que los principios prácticos de uno sean objetivos). Finalmente, si asumo que soy un yo, he de considerar que las razones que me llevan a preocuparme por mi futuro yo, me llevan también a preocuparme por otros “yoes”, pues en ambos casos he de considerar los intereses de entidades que no están presentes en mi conciencia aquí y ahora, en el momento en que tengo que tomar mis decisiones⁴⁰.

Algunos podrían decir que en este razonamiento hay una asimetría pues en el primer caso soy sólo yo el beneficiario de mis acciones, y en el segundo lo son los otros, pero ya hemos visto que esta opción egoísta plantea graves inconsistencias (aunque, curiosamente, asume la tesis de la neutralidad temporal, no puede apoyar, por “razones” egoístas, la de la neutralidad interpersonal). Además, se trata de una objeción poco relevante si, como pretendo, lo que se quiere demostrar no es que, de hecho, la gente asume sus deberes para con otros sino que tiene la capacidad de hacerlo, y que tal capacidad se deriva directamente de su racionalidad (recordemos que ya hemos rechazado la teoría instrumental de la racionalidad y hemos planteado los problemas racionales que padecen las personalidades fragmentadas). O sea, que la adopción de un punto de vista impersonal e imparcial es un ejercicio inevitable en cualquier sujeto racional, desde el momento mismo en que acepta sus propias metas y razones, y les asigna un valor para él. Como señala Gewirth, todo agente realiza juicios evaluativos sobre la bondad de sus propósitos y sobre las condiciones en las que estos se apoyan, por lo que no puede coherente-

³⁹ En *La posibilidad del altruismo* (1970), *cit.*, p. 154. Como señala acertadamente M. CRUZ en el prólogo al libro de D. PARFIT, *Persona, racionalidad y tiempo*, *cit.*, p. 24, “la cuestión que hay que plantearse no es tanto si el yo importa o no, como, más bien, a quién le importa el yo”.

⁴⁰ T. NAGEL en *La posibilidad del altruismo* (1970), *cit.*, pp. 89 y ss.

mente hacer reclamos para sí mismo que niegue a los que están en las mismas circunstancias que él⁴¹.

Nótese que a lo largo de toda la exposición, y también aquí, hemos asumido el autointerés como punto de partida pero esto no excluye la posibilidad de plantear los fines propios desde un punto de vista imparcial –no estrictamente subjetivo, como exige el egoísmo. Es decir, que, en principio, es posible saltar desde esquemas prudenciales, al ámbito de la moralidad, si bien ello requiere asumir la tesis de la neutralidad temporal (como hace la teoría instrumental, aunque rechazando sus presupuestos irracionales) para pasar desde ahí a la perspectiva interpersonal (como pretende hacer el sujeto fragmentado que no asume la tesis de la neutralidad temporal). Aun con todo, este esquema sigue planteando algunas preguntas en el terreno de la moralidad.

A mi juicio, el problema más relevante que plantean estas posiciones es que no parecen ir más allá de la reflexión individual que, aún definida a partir de la satisfacción de ciertas exigencias de imparcialidad o universalidad, se presenta como la vía más aceptable de acceso a la verdad moral (en materia moral cada uno es su propia autoridad epistémica). Por supuesto, esto no significa que el intercambio de opiniones con los otros o la decisión colectiva

⁴¹ A. GEWIRTH, *Reason and Morality*, University of Chicago Press, Chicago, 1978, p. 48. Algunos autores, como G.E. Moore, K. Baier o R.M. Hare, han planteado de este modo el “argumento de la cuestión abierta” que, en realidad, se vuelve contra una definición de lo “bueno” en términos egoístas, dado que sostiene que el bien de algo no puede ser privado, del mismo modo que ninguna cosa puede existir privadamente o para una sola persona. No me resisto a reproducir aquí las palabras de MOORE, a este respecto, en *Principia Ethica* (1903), *cit.*, p. 179: “[...] cuando me refiero a una cosa como “mi propio bien”, todo lo que puedo dar a entender es que algo que será exclusivamente mío [...] es también *bueno absolutamente*, o, más bien que mi posesión de eso es *bueno absolutamente*. *Lo bueno* de eso no puede ser, en ningún sentido posible, “privado” o pertenecerme, tal como tampoco puede *existir* una cosa privadamente o sólo *para* una persona. La única razón que puedo tener para tender a “mi propio bien” es que es *bueno absolutamente* que lo que denomino así me pertenezca, es *bueno absolutamente* que *posea* algo que, si lo tengo, no puedan tenerlo los demás. Pero si es *bueno absolutamente* que *posea* algo que, poseyéndolo, no puedan poseer los otros, y si es *bueno absolutamente* que lo *posea*, entonces cualquiera tiene tanta razón como yo en aspirar a poseerlo como lo poseo yo mismo”. De hecho, para BAIER, “[...] La auténtica razón de ser de una moralidad es producir razones que se impongan a las razones del autointerés” (en *The Moral Point of View: A Rational Basis of Ethics*, *cit.*, p. 309). Véase, también: R.H. HARE, *Moral Thinking: its levels, methods and point* (1981), Oxford Clarendon Press, Oxford, 1992, capítulos 5-7 y, más contemporáneamente, Th. NAGEL, *La última palabra. La razón ante el relativismo y el subjetivismo* (1997), *cit.*, capítulos 2 y 6.

carezcan por completo de relevancia moral, pero, desde luego, tienen un valor auxiliar: “[...] obedecer a otro siempre implica seguir nuestras propias razones, sean éstas morales o prudenciales”⁴²; o sea, sólo la reflexión individual efectiva puede determinar el resultado de un procedimiento colectivo ideal (sometido a determinadas restricciones formales).

Sin embargo, a mi juicio, la exigencia de imparcialidad puede detectarse únicamente de forma cooperativa. Como diría Nino, “la discusión y la decisión intersubjetiva es el procedimiento más confiable de acceso a la verdad moral” porque, aunque no puede excluirse la importancia de la reflexión individual en este campo, resulta, sin duda, menos confiable. Está claro que el individuo ha de salvar siempre la dificultad que supone representarse fielmente los intereses de los demás para lograr adoptar, finalmente, un punto de vista imparcial⁴³. Y aunque pudiera lograrlo, persisten todavía las dudas que suscita ese yo kantiano, ese yo genuino, ese sujeto moral completamente despojado de sus arbitrarias particularidades contingentes. No se me interprete mal. Me parece obvio que los presupuestos formales, como el de imparcialidad, son decisivos para la validez de los principios morales, pero quizá sea conveniente, siguiendo a Habermas, por ejemplo, presentarlos más que como presupuestos formales del razonamiento moral monológico, como reglas de una práctica social del discurso intersubjetivo. Desde luego, esto sigue sin salvar todos los problemas en el ámbito de la moralidad. De hecho, el modelo habermasiano ha sido vulnerable a un buen número de críticas, algunas de las cuales han obligado al autor a revisarse (quizá de forma todavía insatisfactoria). Sin embargo, creo que, debidamente corregido, este esquema ofrece una visión más comprensiva de la racionalidad, y, sin renunciar al autointerés como punto de partida, facilita la construcción de un código moral más aceptable.

5. EL CONSTRUCTIVISMO DIALÓGICO HABERMASIANO. LA RACIONALIDAD COMUNICATIVA COMO FUENTE DE MORALIDAD

Como es bien sabido, en Habermas, la necesidad de pasar de una razón instrumental a una razón dialógica viene avalada por la ética comunicativa.

⁴² C.S. NINO, “Constructivismo epistemológico: entre Rawls y Habermas”, *Doxa*, núm. 5, 1987, pp. 87-105, p. 92.

⁴³ C.S. NINO en “Constructivismo epistemológico: entre Rawls y Habermas”, *cit.*, p. 99 y p. 105 y en *El constructivismo ético*, Centro de Estudios Constitucionales, Madrid, 1989, pp. 105-110.

Aunque es obvio que no sólo ha sido este autor quien ha pretendido armonizar la acción teleológicamente racional, orientada al éxito, y la racionalidad comunicativa, orientada hacia la comprensión intersubjetiva⁴⁴, es cierto que ha sido su propuesta la que ha alcanzado una mayor difusión en nuestros días y la que sigue todavía en proceso de discusión.

Independientemente de que pueden usarse diferentes terminologías, cabe afirmar que en la teoría habermasiana la acción estratégica se rige por reglas técnicas y viene sancionada por su éxito o su fracaso. La acción comunicativa (la interacción en sentido estricto) se rige por normas consensuales que expresan expectativas recíprocas, se funda en la intersubjetividad del entendimiento mutuo, y viene asegurada por el reconocimiento general de obligaciones. Parecería, como afirma T. McCarthy, que la primera no fuera una acción social o no se desarrollara en un contexto social, pero no es esto lo que se deduce de esta teoría, pues se trata de una acción ligada a normas consensuales y tiene lugar en el plano de la intersubjetividad (los actores son capaces de diseñar y seguir sus propias estrategias). Por consiguiente, la distinción entre ambos tipos de acción se apoyaría fundamentalmente en los asuntos que se tematizan en el desarrollo de cada una, en su carácter monológico o dialógico, o en el hecho de que en la primera los intereses individuales predominan sobre consideraciones de reciprocidad, y en la segunda lo decisivo es una reciprocidad basada en el entendimiento mutuo⁴⁵.

Pues bien, si he considerado aquí el estudio de este modelo es porque creo que gran parte de su fuerza reside en el hecho de haber conseguido (con los matices que se quieran) combinar la mejor versión de ambas acciones, ubicando la ética deontológica en la comunicación lingüística, y elaborando, de este modo, un código moral que no se desentiende de los intereses individuales. Y ello aunque, como veremos, el logro del experimento puede ser limitado.

En efecto, Habermas parte de la teoría consensual de la verdad adoptando una postura crítica tanto frente al puro subjetivismo, que reduce las nor-

⁴⁴ Algunos integrantes de la Escuela de Francfort se movieron también por este interés. Así, Horkheimer hablaba de razón objetiva y subjetiva. Esta última era, en la terminología de Adorno, razón identificadora y en la de Marcuse, razón unidimensional. Finalmente, han sido Habermas y Apel quienes han traducido y revisado estos términos, para hablar de razón instrumental/estratégica y razón comunicativa.

⁴⁵ Así lo entiende T. MCCARTHY en *La teoría crítica de J. Habermas* (1978), trad. M. Jiménez Redondo, Tecnos, Madrid, 1998, pp. 46-50.

mas a decisiones, como frente a la ética axiológica objetiva, que concibe las relaciones axiológicas desligadas de su nexo vital real para situarlas en un ámbito ideal del ser que trasciende la experiencia. Lo que Habermas propone es un tipo de intersubjetivismo destinado a explicar y fundamentar consensualmente la verdad de los argumentos y la corrección de las normas (manifestaciones de la praxis comunicativa), considerando que la conciencia individual no se forma en la relación de un individuo con un “mundo exterior” que se le presenta como “objeto”, sino a través de su interacción comunicativa con otros sujetos conscientes.

Desde la ética comunicativa la fundamentación de nuestras normas y principios éticos se apoya en acuerdos intersubjetivamente adoptados mediante el ejercicio de la discusión racional y en el seno de una comunidad de comunicación. De esta manera, una norma sólo es correcta si puede ser el resultado de un procedimiento en el que se hayan respetado las reglas del diálogo, por lo que atender a las reglas procesales que son subyacentes al discurso moral, es decir, a las reglas que la formulación de juicios morales debe satisfacer, resulta imprescindible si se quiere determinar cuáles son los intereses que comparten todos los miembros de una comunidad de hablantes; o sea, el modo en el que el universalismo dialógico puede alcanzarse.

Muy someramente, podemos recordar que las citadas reglas o condiciones pre-requeridas en la comunicación intersubjetiva se relacionan directamente con las *pretensiones de validez* que están presentes en cada acto de habla: pretendemos que lo que decimos es verdadero (pretensión de verdad), que se adecua a las normas morales que rigen este tipo de interacciones (pretensión de rectitud normativa), y que somos sinceros al decirlo (pretensión de veracidad)⁴⁶. De modo que, por lo que aquí interesa, los interlocutores sólo pueden afirmar aquello en lo que verdaderamente creen (se excluye la posibilidad de la automanipulación) y han de reconocerse recíprocamente una igual competencia para el diálogo, dado que no pue-

⁴⁶ J. HABERMAS, *Teoría de la acción comunicativa: complementos y estudios previos*, trad. M. Jiménez Redondo, Cátedra, Madrid, 1984, pp. 121-130. En *Conciencia moral y acción comunicativa*, trad. R. García Cotarelo, Península, Barcelona, 1991, p. 78, dice el autor: “En el caso de las pretensiones de verdad y rectitud, el hablante puede dar cumplimiento a su garantía por medios discursivos, esto es, aduciendo razones; en el caso de las pretensiones de sinceridad, mediante un comportamiento congruente (el hecho de que alguien piense de verdad lo que dice, es algo que sólo puede demostrarse mediante la congruencia de sus actos y no acumulando motivos)”.

den querer para sí mismos lo que niegan a los otros (adoptan un punto de vista imparcial)⁴⁷.

Por un imperativo de brevedad, no voy a detenerme aquí. Me basta ahora con subrayar que en esta construcción la única garantía de la objetividad epistémica de nuestras creencias es precisamente su intersubjetividad, y que alcanzar tal objetividad exige un ejercicio dialógico, más que monológico, de la racionalidad (de ahí que la ética habermasiana sea simultáneamente procedimental, universalista y deontológica). De hecho, es este ejercicio dialógico el que da lugar a la racionalidad comunicativa; una racionalidad que permite a los sujetos plantear sus particulares definiciones e interpretaciones del mundo con el fin de alcanzar un entendimiento, y que requiere, esto es importante, que las pretensiones de validez vengan racionalmente motivadas (puedan cuestionarse).

Hablamos, pues, de una acción comunicativa fuerte dado que el entendimiento entre estos sujetos se extiende a razones normativas compartidas que vinculan sus voluntades con independencia de sus propias preferencias. Por el contrario, cuando la acción comunicativa es débil, el entendimiento se refiere a "hechos y a razones relativas al actor, que sostiene expresiones de voluntad unilaterales"⁴⁸, y las pretensiones de validez que guían a los sujetos son sólo las de verdad o veracidad (no las pretensiones de corrección intersubjetivamente reconocidas). Aquí el entendimiento significa solamente que el oyente entiende el contenido de la declaración de intención o del imperativo del hablante, y no duda de su seriedad, ni de su viabilidad, pero no hay expectativa recíproca de guiarse por normas y valores comunes, ni de aceptar obligaciones mutuas⁴⁹.

Aún a riesgo de subrayar lo obvio, hay que notar aquí, que una acción comunicativa débil no es una acción estratégica. En esta última también faltan los contextos normativos compartidos y las correspondientes pretensio-

⁴⁷ Son las que menciona J. HABERMAS en *Teoría de la acción comunicativa: complementos y estudios previos*, cit., pp. 150-158 y en *Teoría de la acción comunicativa* (1981), segundo vol., trad. M. Jiménez Redondo, Taurus, Madrid, 1992, pp. 211 y ss.

⁴⁸ J. HABERMAS, *Verdad y justificación*, trad. P. Fabra y L. Díez, Trotta, Madrid, 2002, p. 117. Habermas distingue ahí entre el distinto uso que se le puede dar a las expresiones lingüísticas según se orienten o no a la comunicación, y, por lo que hace su uso comunicativo, según la intención sea alcanzar un acuerdo o lograr un auténtico entendimiento (acción comunicativa débil y fuerte, respectivamente). Idem, pp. 109-116.

⁴⁹ *Ibidem*, pp. 118.

nes de corrección normativa, pero no falta sólo eso. En la acción estratégica las pretensiones de verdad y veracidad no se orientan a la motivación racional del oyente sino a que el destinatario extraiga sus propias consecuencias de aquello que el hablante le da a entender. Evidentemente, ambos comparten un saber lingüístico, pero lo relevante es que cada uno se guía por sus preferencias y finalidades, y que no transforman tales preferencias en pretensiones de validez susceptibles de ser reconocidas intersubjetivamente.

En definitiva, sólo hay ejercicio dialógico y acción comunicativa en sentido fuerte cuando las pretensiones de validez se cuestionan públicamente y se hace necesario defenderlas y criticarlas mediante *argumentos* que busquen convencer al resto de los participantes en el diálogo. Es decir, cuando se pone en duda la verdad de enunciados teóricos o la corrección de enunciados prácticos, y se ha de recurrir al *discurso* a fin de restaurar un consenso que permita coordinar nuevamente la acción. Mediante el *discurso* es posible alcanzar una fundamentación racional de las normas, pero sólo cuando éste se desarrolla en un contexto lingüístico regido por una pretensión de rectitud normativa o una pretensión de corrección, puede garantizarse que el consenso logrado muestra la solución racional a los problemas, que no es sino aquella que se alcanza por la fuerza del mejor argumento⁵⁰.

En otras palabras, las pretensiones de validez han de pasar por delante de las meras pretensiones de poder (de las relaciones de poder, manipulación o imposición) para que los potenciales de racionalidad se realicen en la práctica y, de este modo, sea posible alcanzar un consenso en torno a algo que todos puedan querer y que merezca la aceptación y la adhesión general (consenso contrafáctico)⁵¹. En esta construcción una norma sólo es válida en la medida en que puede ganar el asentimiento de todos los afectados; es decir, es válida en la medida en la que es imparcial y ha superado el test de su universalización⁵².

De manera que es la pretensión de rectitud normativa, la expectativa mutua de regirse por valores y normas comunes, que sólo surge en la diná-

⁵⁰ Puede consultarse al respecto, J. HABERMAS: *Teoría de la acción comunicativa* (1981), *cit.*, pp. 43-69.

⁵¹ Véase *Conciencia moral y acción comunicativa, cit.*, pp. 93 y 116.

⁵² En *Conciencia moral y acción comunicativa, cit.*, p. 85, J. HABERMAS, afirma: "Únicamente es imparcial la situación desde la cual son susceptibles de universalización precisamente aquellas normas que, al incorporar de modo manifiesto un interés común a todas las personas afectadas, pueden contar con una aprobación general, así como conseguir un reconocimiento intersubjetivo".

mica de un diálogo real, la que nos obliga al cuestionamiento de nuestras posiciones, la que da lugar a la acción comunicativa fuerte y la que evita que sean las pretensiones de poder las que finalmente se impongan. Y es por esta razón por la que el elemento fáctico en esta construcción no puede perder terreno frente a sus elementos ideales.

Vista esta descripción, podría decirse que la situación ideal habermasiana no se presenta como un dato empírico en sentido estricto pero tampoco como una mera abstracción formal: implica una anticipación por parte de todo el que de hecho entra en un proceso discursivo, de que puede llegarse a un consenso racional (pretensión de corrección). La anticipación de una situación ideal de comunicación supone aceptar el principio de universalización y el compromiso de favorecer la puesta en práctica de las condiciones procedimentales requeridas por una comunicación libre, y este camino lo ejerce la comunidad real de interesados de forma dialógica⁵³. De esta forma se logra soslayar el riesgo de incurrir en una “falacia sustitucionalista” con la que al indagar qué es lo que elegiría la gente si pudiera prescindir de sus intereses, valores, compromisos y “conceptos del bien”, respondiéramos basándonos en nuestros valores, compromisos y conceptos del bien (el paternalismo filosófico al que antes aludí). Y es que, como ya he señalado, para Habermas, la discusión y la decisión colectivas no tienen solamente un valor auxiliar sino que, sometidas a determinadas restricciones formales, se convierten en el procedimiento más confiable de acceso a la verdad moral, lo cual no excluye, una vez más, que por vía de la reflexión individual alguien pueda acceder al conocimiento de soluciones correctas⁵⁴.

⁵³ *Idem.*, pp. 87-88. Habermas se distancia de esta manera de algunas posiciones monológicas que, en su opinión, no pueden mantenerse en una sociedad plural donde conviven distintas visiones del mundo, y que, además, parecerían blindadas frente a revisiones resultantes de futuras experiencias y procesos de aprendizaje (así, frente a Rawls, en “Reconciliación mediante el uso público de la razón”, J. HABERMAS/J. RAWLS, *Debate sobre el liberalismo político*, trad. G. Vilar Roca, Paidós, Barcelona, 1998, p. 52 y p. 54).

⁵⁴ De este modo, al exigir que todos los contenidos, incluso los que afectan a normas de acción fundamentales, se hagan dependientes de discursos reales, Habermas quiere escapar a la crítica formulada contra la tesis que J. RAWLS mantiene en su *Teoría de la justicia* (1971) (trad. M.D. González, Fondo de Cultura Económica, México, 1993), según la cual el postulado de imparcialidad se encuentra satisfecho cuando aquel que estipula el juicio moral se sitúa ficticiamente en la posición de cada uno de los afectados. Las diferencias entre el modelo rawlsiano y habermasiano en este y otros sentidos, las ha señalado el propio HABERMAS en *Aclaraciones a la ética del discurso*, trad. J. Mardomingo, Trotta, Madrid, 2000 pp. 127 y ss.

Pero, vayamos a la segunda cuestión a las que nos referimos al final del apartado anterior, ¿seguimos frente a un sujeto moral desnudo, completamente despojado de sus particularidades contingentes? ¿Es cierto, en definitiva, que esta posición logra situarse entre la abstracción y el dato empírico? Precisamente, por ser esta su pretensión, la teoría habermasiana ha sufrido el embate de las críticas de uno y otro lado. Se le ha criticado tanto su formalismo e, incluso, su hiperracionalismo⁵⁵, como su conexión soterrada con determinados factores culturales, sociales y personales.

Por supuesto, Habermas diría que su ética discursiva es tan inmanente como trascendente, porque no puede encontrarse más allá de los juegos lingüísticos concretos, de los contextos en los que los individuos reales se hallan insertos, pero simultáneamente nos ofrece criterios para criticar nuestras actividades; en ella se toma como punto de referencia la situación del consenso fáctico aunque enmarcado en condiciones trascendentales que aseguran una situación ideal de discurso. Por mor de la pretensión de corrección, todo discurso supone contrafácticamente la meta del acuerdo y, por lo tanto, la existencia de una comunidad ideal; es decir, todo aquel que pretende justificar sus argumentos en el seno de una comunidad real presupone la existencia de una

⁵⁵ Ahí está, por ejemplo, la crítica de A. HELLER, para quien la comunidad ideal habermasiana se construye a partir del dato prácticamente exclusivo de la racionalidad del hombre y de su capacidad lógica de argumentación. “Los hombres en quienes se sostiene este ideal no son hombres enteros. Pues carecen de cuerpo, de sentimientos y ni siquiera tienen relaciones humanas. La relación entre ellos estriba únicamente en la discusión del valor. No es preciso que sean hombres; de la misma forma podría tratarse de ángeles. Pero no destinamos a los ángeles nuestra utopía radical. El hombre es indudablemente un ser racional, pero no es sólo eso” (en *Por una filosofía radical*, trad. J.F. Ivars, El Viejo Topo, Barcelona, 1980, p. 127). La misma idea viene a expresar la autora en *Más allá de la justicia*, trad. J. Vigil, Crítica, Barcelona, 1990, pp. 297 y ss. Tampoco han faltado quienes, como Rorty, mantienen que el momento trascendente de validez universal que sostiene Habermas es una jactancia vacía y que también lo es su intento de conectar facticidad y validez mediante la conexión de dos diferentes usos del lenguaje: el estratégico y el comunicativo. De hecho, según Rorty, tal diferenciación no puede darse porque todas las prácticas están sometidas a convenciones sociales y son superadas, únicamente, por otras prácticas sociales; no hay pues nada parecido a una razón comunicativa orientada a la verdad (véase: R. RORTY, *El pragmatismo, una versión. Antiautoritarismo en epistemología y ética*, trad. J. Verés Gifra, Ariel, Barcelona, 2000, pp. 92-100). “Mi problema se intensifica al formularme la pregunta de si mis pretensiones de verdad “trascienden mi contexto cultural local”. Como no puedo comprender qué significa aquí “trascendencia”, no tengo claro si lo hacen o no. Y es que ni siquiera comprendo qué sentido tiene decir que mi afirmación “tiene una pretensión de verdad”” (*op. cit.*, p. 102).

comunidad ideal a través de la cual se determina la corrección de tales argumentos⁵⁶. De este modo, aunque se distingue el mundo fáctico de la eticidad y el mundo contrafáctico de la moralidad, ambos pueden reconciliarse⁵⁷. De hecho, la comunidad ideal de comunicación sería el fundamento del principio de universalización que patrocina los acuerdos reales, y también la última garantía de su racionalidad, esto es, “la razón última a la que acudir cuando nos preguntamos por qué debemos hacer lo que debemos”⁵⁸. Y he aquí la vinculación entre racionalidad comunicativa y discurso moral.

Pues bien, de la multitud de críticas que se le han lanzado a este modelo, las que más interesarían a este artículo serían las que lo acusan de excesivo racionalismo, porque si fueran atinadas, si a Habermas le fallara el elemento de la facticidad, no habría forma de distinguir su apuesta teórica del recurso monológico del observador imparcial. Para distanciar a Habermas de este recurso tendríamos que evaluar, seguramente, en qué medida ha logrado suavizar el elemento trascendental de la razón kantiana o situar su comunidad de comunicación en un mundo social⁵⁹. Lamentablemente, no puedo adentrarme en profundidad en un asunto como este.

Diremos, más toscamente, que, en Habermas, el yo no es el fruto de un acto aislado de autorreflexión sino de un proceso de formación en el seno de un

⁵⁶ Siguiendo a K.O. APEL, podría decirse que este principio ético-normativo no puede ser discutido sin contradicción pragmática del que argumenta, sin incurrir en una inconsistencia performativa (“¿Es la ética de la comunidad ideal de comunicación una utopía?”, *La transformación de la filosofía*, vol. II, trad. A. Cortina, J. Chamorro y J. Conill, Taurus, Madrid, 1985, p. 205). Esta metanorma puede explicarse del siguiente modo: con el acto de argumentar, el argumentante ya ha reconocido que la razón es práctica, es decir, que sus pretensiones de validez y de verdad pueden y deben ser satisfechas con argumentos (K.O. APEL: “Necesidad, dificultad y posibilidad de una fundamentación filosófica de la ética en la época de la ciencia”, en *Estudios Éticos*, trad. de C. Santiago y revisión de E. Garzón Valdés y R. Zimmerling, Alfa, Barcelona, 1986, p. 161).

⁵⁷ También McCARTHY asume la virtualidad de la teoría habermasiana para mediar entre sendas opuestas que no han conseguido superar sus propias aporías (en *La teoría crítica de J. Habermas* (1978), *cit.*, p. 456). El propio Habermas lo reafirma en *Acción comunicativa y razón sin trascendencia*, trad. P. Fabra, Paidós, Barcelona, 2002, p. 18, aunque, como se verá, ha revisado algunos de sus presupuestos.

⁵⁸ Así lo señala J. MUGUERZA en “Más allá del contrato social (Venturas y desventuras de la ética comunicativa)”, *Desde la perplejidad (Ensayos sobre la ética, la razón y el diálogo)*, Fondo de Cultura Económica, México, 1995, p. 305.

⁵⁹ Parece que Habermas ha admitido la necesidad de mitigar esta elemento para pasar a un supuesto “pragmatismo de corte kantiano” que desarrolla en *Verdad y justificación* (en el Prólogo a su *Acción comunicativa y razón sin trascendencia*, ya citada).

complejo de interacciones; de un proceso dialógico de socialización. En su esquema, la constitución de un mundo de la experiencia sensible es el resultado de una interacción sistemática entre receptividad sensible, acción y representación lingüística, por lo que es posible pasar desde lo que cada cual puede querer sin contradicción que se convierta en una ley universal (Kant) a lo que todos pueden acordar que se convierta en una norma universal. En la primera posibilidad, las inclinaciones particulares y contingentes quedan excluidas, mientras que en la segunda se trata de llegar a un acuerdo racional acerca de ellas. Por consiguiente, lo que parece exigir Habermas no es la supresión de las inclinaciones personales sino su racionalización, su formación por medio de una comunicación no distorsionada⁶⁰. De este modo, pretende salvar la distancia que pudiera existir entre la abstracción y la idealización (con la que se privilegiaría, de forma encubierta, un determinado tipo moral), o atender al punto de vista del “otro concreto”, en la línea que le ha señalado, por ejemplo, Seyla Benhabib⁶¹.

Sin embargo, aunque estas han sido sus pretensiones no está claro que Habermas haya conseguido aterrizar lo suficiente su construcción como para que, todavía desde paradigmas racionalistas, no haya razones para la crítica. Es demasiado tentador señalar que, de entre todas ellas, persiste la que le acusa de mantener la clásica separación entre las cuestiones de justicia y las referentes a la “vida buena”; una separación que se conecta con la que divide el ámbito público y el privado, y cuya problematización es vital para los objetivos emancipatorios del feminismo. Esta crítica es, además, más hiriente si concordamos en que el compromiso con el procedimiento discursivo universalista como procedimiento de acceso a la verdad moral, no conlleva una delimitación sesgada de lo moralmente relevante y, por tanto, no debe excluir *a priori* el debate acerca de lo que es de suyo bueno⁶². En fin, es-

⁶⁰ Idem, p. 39 y p. 60.

⁶¹ Es imposible mencionar siquiera a la mayoría de las autoras feministas que se han opuesto al contractualismo como base de una teoría de la justicia, pero creo que, por lo que hace a Habermas, y adoptando todavía una perspectiva racional, hay que destacar a S. BENHABIB en *El ser y el otro en la ética contemporánea: feminismo, comunitarismo y postmodernismo* (1992), cit., pp. 171 y ss.

⁶² Este enfoque, como señala C. AMORÓS, tiene la virtualidad de dotar al feminismo de recursos argumentativos para polemizar con las concepciones del bien en las que se da por buena la subordinación de las mujeres, y que son incompatibles con su emancipación (entendida como una exigencia de justicia). Una vez más, el universalismo interactivo de Benhabib es especialmente lúcido y aprovechable como instancia crítica. Véase: *Tiempo de feminismo. Sobre feminismo, proyecto ilustrado y postmodernidad*, Ediciones Cátedra, Universitat de València, Instituto de la Mujer. Colección Feminismos, Madrid, 2000, pp. 401-405.

to es, sin duda, materia para varios artículos más, pero no quería dejar de señalarlo.

Para concluir, puede decirse que el concepto de racionalidad estratégica del que partíamos no sólo resulta insuficiente como explicación de lo que hacemos, sino que parece también poco útil para indicar lo que no deberíamos hacer. El sujeto egoísta en el que se apoya sólo puede plantearse fines subjetivos desde un punto de vista subjetivo y es incapaz de optar por el principio no egoísta de decisión (automanipularse) que necesita para lograrlos. Nada de esto significa que tengamos que renunciar al autointerés como punto de partida para hablar de racionalidad, sino que es precisamente tal presupuesto el que nos obliga a definir nuestros fines desde un punto de vista imparcial. En otras palabras, el sujeto racional no es el que incorpora fines estrictamente personales a su razonamiento o el que afirma sin más sus intereses propios, sino aquel que logra comprometerse con ciertos principios normativos. Es en este ejercicio racional en el que nos reconocemos como seres morales, y no en el de una razón subjetiva que se ocupa sólo del mecanismo del pensamiento (la capacidad de clasificación, inferencia y deducción) y de los medios apropiados para la obtención de fines que sólo sirven a la autoconservación.

Es obvio que esta posición podría sintonizar bien (sin llegar a coincidir) con las críticas que desde hace años se han venido lanzando contra la subjetivización de la razón o contra la hegemonía de la razón instrumental que, en palabras de Horkheimer, lleva a la derrota y al eclipse de la razón⁶³. Recordemos que, según este autor, la irracionalidad racionalizada (racional en los medios e irracional en los fines) supone el sometimiento total de la razón a la realidad social, a la utilidad y al éxito⁶⁴. Sin embargo, rechazar la subjetivi-

⁶³ M. HORKHEIMER, *Crítica de la razón instrumental* (1947), trad. J. Muñoz, Trotta, Madrid, 2002, pp. 10-11 y 47 y ss. En esta obra, Horkheimer insiste en los puntos en los que insiste también en *Dialéctica de la ilustración. Fragmentos filosóficos* (trad. J.J. Sánchez, Trotta, Madrid, 1997), escrita junto con T.W. Adorno en 1947 (al respecto las páginas 251 y ss.). Ambos consideran que la perversión y el mal uso de la razón se encontraban, desde los inicios, en su constitución moderna, si bien su índole dual (emancipación/ dominación) es anterior a la Ilustración. De esta manera, rechazan por unilateral y lineal, la interpretación progresista de la historia de la razón.

⁶⁴ M. HORKHEIMER, *Crítica de la razón instrumental* (1947), *cit.*, pp. 75 y ss. Sobre esta idea, resulta interesante consultar la obra de C. THIEBAUT, *Vindicación del ciudadano. Un sujeto reflexivo en una sociedad compleja*, Paidós, Barcelona, 1998 y la de Z. BAUMAN, *La sociedad individualizada*, trad. M. Condor, Cátedra, Madrid, 2001.

vización de la razón no tiene que provocarnos la nostalgia de una razón dogmático-sustantiva (como diría Rorty)⁶⁵, ni tampoco ha de llevarnos a romper con la razón técnica, tal como hiciera Marcuse, en línea con las primeras tentativas de la escuela de Francfort⁶⁶, sino que sólo nos anima a mantener una teoría más comprensiva de la racionalidad. Como hemos visto, esta teoría ha de pasar por una reconceptualización de la razón como razón dialógica, contextualizada y dinámica, y en la que no se pierda de vista que somos sujetos “reales”, con experiencias particulares, pero que tenemos la capacidad de construirnos a nosotros mismos como sujetos racionales.

MARÍA EUGENIA RODRÍGUEZ PALOP

Instituto de Derechos Humanos Bartolomé de las Casas

Universidad Carlos III de Madrid

c/Madrid, 126

Getafe 28903 Madrid

e-mail: merpalop@der-pu.uc3m.es

⁶⁵ R. RORTY en *Contingencia, ironía y solidaridad*, trad. A.E. Sinnott, Paidós, Barcelona, 1991, p. 76 y en “La prioridad de la democracia sobre la filosofía” (1984), *Objetividad, relativismo y verdad*, trad. J. Vigil Rubio, Paidós, Barcelona, 1996, p. 242. No hay que olvidar que para Rorty, la idea misma de verdad es un resto mítico del pensamiento idealista occidental.

⁶⁶ MARCUSE insiste en la ceguera que acompaña a esta visión instrumental de la razón y que la convierte en una razón destructiva. Una razón que no sólo impide la reflexión sobre intereses sociales sino que los reduce a posibles relaciones de control técnico y de dominio. En opinión de Marcuse, el horizonte instrumentalista de la razón abre paso a una sociedad totalitaria a la que sirve como justificación. Véase al respecto, *El hombre unidimensional: ensayo sobre la ideología de la sociedad industrial avanzada* (1964), trad. Antonio Elorza, Ariel, Barcelona, 1990, p. 53.