

Tras el consenso. Sobre el giro epistemológico-político de John Rawls

MIGUEL GIUSTI

Instituto de Investigaciones Filosóficas,
Universidad de Lima

En su última obra, *Political Liberalism*, John Rawls modifica el estatuto epistemológico de su teoría de la justicia, de forma tal que ésta ha de considerarse ahora como una teoría política y ya no como una teoría moral o metafísica. El giro se lleva a cabo con la finalidad de asegurar la «estabilidad» de la teoría, es decir, su capacidad de generar un *overlapping consensus* con todas las doctrinas morales actualmente en conflicto. El tema de nuestro trabajo es el análisis del rol central que ahora desempeña en la teoría este *overlapping consensus*. De él habla Rawls como una meta

por alcanzar, cuando la teoría entra en contacto con las doctrinas rivales; se puede decir, en tal sentido, que *andamos tras el consenso*. Pero este consenso perseguido reposa, en realidad —ésta es nuestra tesis—, sobre un acuerdo o consenso previo, cuya validez debe ser asumida por la teoría en el nuevo marco epistemológico; el consenso nos precede, *nos hallamos tras él*. El análisis de esta ambivalencia del *overlapping consensus* permite mostrar las debilidades epistemológicas del giro anunciado por Rawls.

Todos los caminos de la filosofía política contemporánea parecen conducir a John Rawls. Para seguir sus pasos o para desandarlos. Para emular su propuesta teórica o para hacer de ella el blanco de las críticas, pero teniendo siempre como punto de referencia el modelo de justicia imparcial asociado ya indisolublemente a su nombre. Su teoría de la justicia como *fairness*¹, anunciada en ensayos sistemáticos desde fines de los años cincuenta, contribuyó a dar un giro significativo a la filosofía política anglosajona, pues reactualizaba en su seno planteamientos contractualistas y planteamientos kantianos concernientes a la fundamentación del universalismo moral. Desde entonces, la figura de John Rawls parecía encarnar, casi paradigmáticamente, la concepción moderna de la moral en su versión liberal. Que su figura era tan representativa lo confirma, entre otras cosas, el hecho de que desde los años ochenta se

¹ J. Rawls, *A Theory of Justice*, Oxford, Oxford University Press, 1971. Traducción castellana, con el título *Teoría de la justicia*, de María Dolores González, México, FCE, 1979. Esta obra condensa, en realidad, al igual que lo hará luego *Political Liberalism*, una serie de ensayos anteriores de Rawls, cuyo núcleo sistemático apareció en 1958 bajo el título «Justice as Fairness» [*The Philosophical Review*, LXVII (1958), pp. 177-197].

haya convertido en el principal blanco de las críticas de los autores vinculados al comunitarismo. En el retrato polémico que los comunitaristas hacen del liberalismo, los rasgos definitorios principales son extraídos premeditadamente de la concepción de Rawls². Pero, si parecía así haberse planteado con nitidez una oposición irreconciliable entre ambos modelos en disputa, en sus últimos trabajos John Rawls ha efectuado un importante giro epistemológico que viene a desarticular la polaridad de la discusión, generando desconcierto entre sus detractores y sus defensores.

Todos éstos son, por cierto, desplazamientos conceptuales que se operan primeramente en la tradición anglosajona, más específicamente en la tradición norteamericana, de la filosofía política y que no coinciden necesariamente con la evolución de esta reflexión en otros contextos. En el contexto de la filosofía moral alemana, por ejemplo, ni el neokantianismo rawlsiano ni el retorno comunitarista al *ethos* tendrían por qué ser considerados novedosos, pues sobre tales motivos hay en la tradición alemana una disputa ya antigua, que en las últimas décadas se dio a conocer como el debate entre «moralidad» y «eticidad»³. Desde esta perspectiva, la polémica de Rawls contra el utilitarismo o contra el intuicionismo podría ser vista como una polémica anticuada o, peor aún, como una polémica contra adversarios inexistentes. Pero, como es obvio, otro tanto podría decirse, desde una perspectiva anglosajona, sobre las posiciones conceptuales del debate alemán. En realidad, esto es algo que no debería sorprendernos tanto: que los filósofos norteamericanos combatan el utilitarismo, que los alemanes combatan el relativismo o que los franceses combatan el racionalismo (es decir, que combatan a los dos anteriores) parece ser una demostración inadvertida e irónica de la influencia de los contextos sobre los debates acerca del contextualismo moral.

Pero volvamos a Rawls. El giro que éste venía anunciando a través de lecciones y breves ensayos en los últimos quince años parece haber hallado su culminación en el nuevo *opus magnum* publicado en 1993: *Political Liberalism*. Pero ¿en qué consiste este giro? ¿Qué ha cambiado de la originaria teoría de la justicia y por qué fue necesario ese cambio? Si nos atenemos a lo que Rawls mismo afirma, todos los intentos parciales y dispersos de reformulación de sus planteamientos iniciales habrían tenido por finalidad afrontar una inconsistencia central de su teoría, a saber: su incapacidad para resolver el problema de la «estabilidad»⁴. «Estabilidad» significa aquí la capacidad de una teoría

² El caso más manifiesto es el de Michael Sandel, *Liberalism and the Limits of Justice*, Cambridge, Cambridge University Press, 1982. Pero otro tanto podría fácilmente detectarse, en forma más o menos explícita, en los trabajos de Alasdair MacIntyre o Charles Taylor.

³ Me remito a dos trabajos que he escrito anteriormente sobre el sentido y el desarrollo de este debate: M. Giusti, «Lassen sich Moralität und Sittlichkeit miteinander vermitteln?», *Philosophische Rundschau*, 38 (1991), pp. 14-47, y «Moralidad o eticidad. Una vieja disputa filosófica», *Hueso Húmero* (Lima), 28 (1991), pp. 54-75 [aparecido igualmente en *Estudios de Filosofía* (Universidad de Antioquia), 5 (1992), pp. 49-64].

⁴ J. Rawls, *Political Liberalism*, Nueva York, Columbia University Press, 1993, pp. xv-xvi.

de hallar para sus hipótesis un fundamento tal que pueda contar perdurablemente con el asentimiento de todos, sin que nadie deba verse por ello obligado a renunciar a su propia posición. Una teoría será, pues, «estable» en la medida en que logre generar un consenso en el que converjan o coincidan, al menos en lo esencial, todas las posiciones a las que ella (la teoría) se dirige. Habrá de ser, en tal sentido, un *overlapping consensus*, un consenso en el que las partes involucradas llegan a superponerse o a entrecruzarse unas con otras. (A falta de una mejor solución, y para evitar la traducción española de «consenso por solapamiento», muy desafortunada fuera de la península, lo llamaré de ahora en adelante un «consenso abarcador».) Si el problema de su obra original era, pues, la «estabilidad», la forma de resolverlo es ahora la búsqueda de un «consenso abarcador».

Pero ¿por qué cree Rawls que *Una teoría de la justicia* no satisfacía esta exigencia? Porque en ella, nos dice, la tesis central era propuesta en forma de una concepción moral con pretensiones absolutas, que se hallaba por ende en rivalidad con otras concepciones morales análogas, en lugar de tenderles algún puente conceptual capaz de motivar un consenso en cuestiones esenciales. «Este problema [de la estabilidad] —escribe Rawls— constituye el escenario en el que se desenvuelven los ensayos recientes a partir de 1980... Creo que todas las diferencias derivan de haber despejado aquella inconsistencia»⁵.

El término «estabilidad» no parece muy afortunado, a decir verdad, para caracterizar la naturaleza del giro mencionado. Por prestarse justamente a malentendidos, ha sido ya objeto de una severa crítica, que le reprocha a Rawls estar adoptando la posición de un político pragmático dispuesto a negociar con los otros la vigencia de su teoría⁶. Para hacer frente a esta objeción, que él rechaza categóricamente, Rawls nos advierte que es preciso diferenciar dos fases en la constitución de su teoría: en una primera fase (o en un primer nivel), ésta debe ser una «concepción autónoma» (*free standing view*), que se legitima por sí sola sin restricciones derivadas de su contrastación con otras; y sólo en una segunda fase (o en un segundo nivel) ella habrá de afrontar el problema de la estabilidad⁷. En otras palabras, la teoría no es negociable porque en el momento mismo de su formulación, es decir, en su primera fase, ella posee ya una justificación adecuada y suficiente. Pero esta diferenciación, lejos de eliminar las ambigüedades de la nueva propuesta, las pone más al descubierto aún, en la medida en que le resta claridad al rol que juega

⁵ *Ibidem*, pp. xvi-xvii.

⁶ Es el caso, entre otros, de J. Raz: «Facing Diversity: The Case of Epistemic Abstinence», *Philosophy and Public Affairs*, 19 (1990), pp. 3-45.

⁷ Véase «The Domain of the Political and Overlapping Consensus» [*New York Law Review*, 64 (1989), pp. 233-255], donde Rawls introduce por primera vez esta diferenciación, pero sobre todo *Political Liberalism*, IV. Ahora divide Rawls las dos primeras partes de su libro siguiendo esta diferenciación, de modo que la primera parte se ocupa de la teoría en tanto *free standing view* (es decir, de su primera fase), mientras que la segunda parte se ocupa de la búsqueda de un «consenso abarcador» (de su segunda fase).

en ella el consenso, así como a las razones por las que éste habría de ser asumido. Explicaré mejor esta dificultad, pues ella servirá de hilo conductor de mi argumentación.

El giro actual de Rawls se explica, veámos, por la necesidad de resolver el problema de la estabilidad (de la falta de estabilidad) de su teoría de la justicia. Pero como la inestabilidad se debe a su incapacidad para generar un acuerdo con otras teorías rivales, la solución consiste en buscar puntos de contacto que permitan la formación de un «consenso abarcador». El consenso del que aquí se habla es un consenso por alcanzar. *Andamos tras él*. No obstante, para que la teoría no se exponga a la objeción de estar negociando su validez, Rawls nos dice que ella deberá primero justificarse autónomamente como concepción. Pero, en tal caso, ya no se entiende muy bien cuál es el valor o el alcance del consenso por alcanzar, dado que la teoría debe disponer, de antemano, no sólo de los recursos para autolegitimarse, sino además de los criterios por los que se habrá de regir luego la generación del consenso; si no contara con ambas cosas, peligraría su autonomía. Ni se entiende tampoco, en consecuencia, con qué finalidad se mantiene aún la división de la teoría en dos fases, ya que los elementos indispensables para la instauración del consenso —de la segunda fase— deben hallarse ya disponibles cuando la teoría entre en contacto con las demás, es decir, deben definirse en su primera fase. La cuestión principal que la teoría debe afrontar es por eso la de definir la naturaleza del consenso previo, aquel que le sirve de fundamento y que le permite aproximarse a las otras con la esperanza de obtener su adhesión. El consenso del que aquí se habla es un consenso que nos precede. *Nos hallamos tras él*.

Con el título de mi trabajo —«Tras el consenso»— quiero hacer, pues, explícita la ambivalencia latente en el giro epistemológico de Rawls. «Tras el consenso» puede querer decir, tal como él se lo propone, ir en su búsqueda, efectuar las modificaciones necesarias para que la propia teoría se halle en condiciones de entablar un diálogo conciliador con expectativas de éxito. De ello me ocuparé en la segunda de las dos partes de este ensayo: «El consenso perseguido». Pero «tras el consenso» puede querer decir también, tal como será sugerido aquí, darlo por presupuesto, apoyarse en un fundamento que permita a la teoría confiar en su propia legitimidad y en la validez de sus pretensiones para obtener el asentimiento ajeno. De ello me ocuparé en la primera parte: «El consenso presupuesto».

I. *El consenso presupuesto*

La noción misma de «estabilidad» encierra ya una idea de lo que la teoría debe ser y dejar de ser. Para ser estable, ella debe empeñarse en hallar fórmulas de acercamiento a otras concepciones. Pero para justificar teóricamente este

cambio de actitud, ella debe modificar su estatuto epistemológico originario, que sólo le permitía presentarse en condición de competidora excluyente. La primera y principal transformación de la teoría de la justicia consiste por eso en considerarla en adelante como una teoría exclusivamente «política», ya no «moral» en sentido estricto, ya no «metafísica»⁸. Restringiendo su atención al dominio estructural de la sociedad y al dominio público, debe diferenciarse de las «doctrinas morales o filosóficas» que tienen pretensiones «más comprensivas» en virtud de alguna concepción específica del bien o de la realidad⁹. «En *Una teoría de la justicia* —escribe Rawls— no se dice nada sobre el contraste existente entre las doctrinas morales y filosóficas comprensivas y las concepciones que se restringen al dominio de lo político»¹⁰. Doctrinas morales comprensivas —precisamente lo que la teoría debe dejar de ser— son aquellas, al decir de Rawls, que, valiéndose de ciertos principios, creen poder dar una respuesta a todas las preguntas concernientes a la vida humana o aquellas que se consideran portadoras de una concepción de la vida buena¹¹. Muchas concepciones filosóficas, morales o religiosas podrían ajustarse a esta caracterización global, sin que ninguna de ellas pueda pretender seriamente tener de su parte argumentos convincentes para probar la validez universal de sus afirmaciones. La inconmensurabilidad de las teorías morales en pugna es para Rawls un hecho, un dato, que constituye el punto de partida de la filosofía política de la actualidad¹².

Por cierto, esta idea de que la filosofía deba hacer frente a un conflicto entre concepciones morales irreconciliables entre sí no es nueva, y algunas de las doctrinas por Rawls llamadas «comprensivas» (por ejemplo, la moral kantiana) hallan su razón de ser justamente en la búsqueda de una solución a aquel dilema. Lo nuevo es, más bien, que también las soluciones filosóficas a este viejo problema son consideradas ahora como puntos de vista contrapuestos que tan sólo reproducen la situación aporética en un nivel más alto de abstracción. Y así como ya una vez se pudo recurrir, en el terreno de la praxis política, al principio de la tolerancia, así también se trata ahora de «aplicar el principio de la tolerancia a la filosofía misma»¹³. En otras palabras:

⁸ Ésta fue la tesis central que dio nombre al ensayo de Rawls: «Justice as Fairness: Political not Metaphysical», *Philosophy and Public Affairs*, 14 (1985), pp. 223-251.

⁹ Cfr. *Political Liberalism*, *op. cit.*, I: 1.

¹⁰ *Ibidem*, *Introduction*, p. xv.

¹¹ Cfr. *ibidem*, II: 3.

¹² Esta idea de la inconmensurabilidad entre puntos de vista morales es en realidad un motivo metodológico recurrente en la filosofía moral moderna, y no es de extrañar que ahora, en el contexto de crítica de la racionalidad, vuelva a adquirir vigencia. Pero tal idea no conduce necesariamente, como Rawls se lo imagina, a fortalecer la posición del observador imparcial. Un buen contraejemplo es Alasdair MacIntyre, quien toma como punto de partida igualmente el *factum* de la inconmensurabilidad (en *After Virtue*), pero para luego justificar la adhesión a una determinada tradición (a una determinada «doctrina moral comprensiva») (en *Whose Justice, Which Rationality*).

¹³ «Justice as Fairness: Political not Metaphysical», *op. cit.*, p. 231.

ya no tiene sentido seguir tratando de demostrar la necesidad de un acuerdo a través de estrategias de argumentación supuestamente irrefutables, pues todas las concepciones particulares, sin coincidir unas con otras, reclaman para sí un análogo e idéntico derecho. Manifiestamente, el método habitual de la confrontación no conduce a satisfacer las expectativas consensuales que en él se depositan y debe ser por eso abandonado. El propósito explícito de la teoría no es ya solucionar, sino *evitar* el conflicto ¹⁴.

Evitar el conflicto significa, para Rawls, abstenerse de un juicio sobre los temas que motivan el desacuerdo, pero abstenerse con el propósito de buscar una base más profunda de comunicación que permita tender puentes de conciliación entre las partes implicadas en la controversia. Pero, como es obvio, a esta base no podrá accederse por los mismos medios que condujeron a la situación de incompatibilidad que se trata ahora de superar. Es preciso efectuar un giro metodológico. Y el giro consiste en remontarse a la tradición de la sociedad democrática, en cuyo seno han ido surgiendo paulatinamente las doctrinas comprensivas que hoy compiten entre sí. «Nosotros consideramos a nuestra cultura política pública —escribe Rawls—, incluidas sus instituciones más importantes y la tradición histórica de sus interpretaciones, como un fondo común de principios e ideas básicas implícitamente reconocidos» ¹⁵. La base de la concordancia ha de buscarse, pues, en este trasfondo sedimentado de la tradición, el cual constituye para Rawls, en igual forma que la existencia de una pluralidad de doctrinas morales rivales, un hecho indiscutible de la cultura política de una sociedad democrática.

Una de las consecuencias del cambio de perspectiva metodológica es que la teoría política debe ahora renunciar a toda pretensión de verdad o, más exactamente: debe *evitar* tal pretensión. Debe hacerlo porque la idea de una verdad moral no deja espacio alguno para la coexistencia de concepciones discrepantes, que tienen todas ellas pretensiones de verdad irreconciliables entre sí. «Es con esta idea de verdad moral —sostiene ahora Rawls— con la que se contrastará la idea de los primeros principios como *razonables*» ¹⁶. «Razonable» es aquí *reasonable*, que no debe confundirse con «racional» *rational*, como veremos en seguida. Por el momento retengamos simplemente que la diferenciación entre lo «razonable» y lo «verdadero» le sirve a Rawls para fijar conceptualmente la toma de distancia frente a los criterios de validación inmanentes a cualquier concepción (frente a sus criterios de *verdad*) y la necesidad de buscar un fundamento de otra naturaleza (de naturaleza *razonable*) para obtener un consenso.

¹⁴ Cfr. *Political Liberalism*, *op. cit.*, I: 2.3.

¹⁵ «Justice as Fairness: Political not Metaphysical», *op. cit.*, p. 228.

¹⁶ J. Rawls, «El constructivismo kantiano en la teoría moral», en J. Rawls, *Justicia como equidad. Materiales para una teoría de la justicia*, traducción de M. A. Rodilla, Madrid, Tecnos, 1986, p. 174.

Con esta forma de proceder, Rawls retoma su vieja propuesta metodológica de un «equilibrio reflexivo» (*reflective equilibrium*)¹⁷ y le otorga ahora un lugar central en la reestructuración epistemológica de su teoría. Ya en *Una teoría de la justicia*, el método del equilibrio reflexivo había sido introducido con la finalidad de reemplazar la definición tradicional de fundamentación, es decir, aquella tributaria del modelo lógico-formal de deducción. En su lugar, debía hacerse valer un tipo de fundamentación más débil, aunque con mayores expectativas de éxito, que consiste en comparar, sopesar («equilibrar»), de un lado, nuestras convicciones morales más inmediatas y, de otro, las presuposiciones que subyacen a tales convicciones, y en buscar formas diversas de interrelación entre ambas instancias hasta dar con un terreno adecuado de concordancia que pueda servir de punto de apoyo para la argumentación posterior. No obstante, esta idea perdía importancia en el conjunto de la obra primera, porque la función principal de fundamentación parecía recaer entonces en la ficción de la «posición originaria». Ahora, en cambio, en el nuevo marco epistemológico de la teoría, al equilibrio reflexivo le corresponde —o debería corresponderle al menos¹⁸— desempeñar aquella función esencial.

Se trata, como decíamos, de un tipo de fundamentación, pero no deductiva. Esto último es fácil de entender, y refuerza además el sentido de la diferenciación entre lo «razonable» y lo «verdadero». Una cadena deductiva de argumentos que parte de premisas primeras puede tener sentido en la búsqueda de la verdad, pero en la búsqueda de un acuerdo práctico (razonable) ella es más bien un obstáculo. «Tenemos que distinguir —escribe Rawls— entre el orden deductivo (*order of deduction*) y el orden sustentador (*order of support*). Los argumentos deductivos prescriben el orden en que han de ser encadenadas las ideas... Pero las ideas que justifican una concepción normativa y nos garantizan que ella es razonable, no necesariamente ocupan un lugar superior en el orden de la deducción. Si clasificamos los principios y las convicciones de acuerdo a qué tan fuertemente hacen de sustento a la doctrina que los anima, veremos que los principios y convicciones que ocupan un lugar superior en este orden sustentador pueden ocupar un lugar inferior en el orden deductivo. La noción de equilibrio reflexivo da a entender que las razones que hacen de sustento pueden ser proveídas por convicciones de cualquier nivel de generalidad»¹⁹. Esta larga cita nos indica con suficiente claridad en qué dirección debe orientarse la fundamentación por contraste con el procedimiento deduc-

¹⁷ Cfr. *Teoría de la justicia*, *op. cit.*, pp. 37 y s. y 68 y ss.

¹⁸ Es preciso señalar, sin embargo, que Rawls no desarrolla la noción de *reflective equilibrium* en su último libro con la precisión que se esperaría, dado el rol central que le corresponde ahora en la reformulación de su teoría. Esto llama la atención sobre todo porque desde muy temprano se entabló una discusión entre los críticos e intérpretes de Rawls en torno al estatuto del equilibrio reflexivo. Cfr., por ejemplo, N. Daniels: «Reflective Equilibrium and Archimedean Points», en *Canadian Journal of Philosophy*, X (1980), pp. 83-103.

¹⁹ *Political Liberalism*, *op. cit.*, p. 242, nota 31.

tivo. Y nos hace caer en la cuenta además que nos hallamos aquí en el nivel más elemental, en la instancia última de justificación, de la teoría. Por eso advierte Rawls que no confundamos este punto de vista con el de las partes en el estado originario ni con el de los ciudadanos en una sociedad bien ordenada. Porque las partes son actores *racionales* (ficticios) en el marco de una construcción heurística y los ciudadanos son sujetos *razonables* (ideales) en el marco de un modelo político-moral, mientras que aquí nos ocupamos más bien de las condiciones que hacen las veces de «orden sustentador» tanto de aquel modelo como de aquella construcción.

Ahora bien, si la fundamentación de la que hablamos implica una búsqueda retrospectiva de un sustrato común generador de consenso, es necesario suponer, si es que el método realmente funciona, que dicha búsqueda llegará en algún momento a un fin, o a un comienzo, es decir, que descubrirá aquel sustrato. ¿Cuál debe ser entonces, podríamos preguntarnos, el comienzo de la teoría de la justicia? La respuesta de Rawls: el comienzo es la idea de la sociedad como un sistema de cooperación imparcial entre personas libres e iguales. Ésta es «la idea matriz fundamental... con la que se hallan sistemáticamente conectadas todas las otras ideas básicas»²⁰ de la teoría. La teoría la asume como una idea enraizada en la tradición de la democracia y hace por eso de ella su punto de partida. «Comenzamos —escribe Rawls— con esta idea, que consideramos implícita en la cultura pública de una sociedad democrática»²¹. Pero ¿en qué consiste, más específicamente, semejante «cooperación imparcial»? Para definir la cooperación, Rawls se vale de un modelo de razón práctica en el que ésta posee ya las propiedades de la intersubjetividad, el intercambio de argumentos y la dimensión pública. Y asume así, indirectamente, que este modelo de razón, al igual que la idea de cooperación a la que se halla asociado, obtiene su legitimidad o su vigencia de la tradición de las sociedades democráticas. En efecto, afirma Rawls que cooperación es la acción guiada «por reglas y procedimientos públicamente reconocidos», y que «condiciones de imparcialidad (de *fairness*)» son aquellas «que todo participante reconocería razonablemente, suponiendo que todos los demás también las reconociesen»²². Ambas nociones parecen, pues, condicionarse recíprocamente, la noción de cooperación (en sociedad) y la noción de razón (intersubjetiva). Su interdependencia se ha hecho más manifiesta en los últimos trabajos de Rawls, en los que la idea de la razón práctica (y de la razón pública) pasan a ocupar un lugar preminente, pero no por ello se ha hecho más clara. Cuando uno desea saber qué significan los conceptos de «cooperación» o de «sistema imparcial» que caracterizan al sustrato consensual originario se ve remitido a un paradigma de racionalidad caracterizado por la

²⁰ *Ibidem*, p. 15.

²¹ *Ibidem*.

²² *Ibidem*, p. 16.

reciprocidad y el intercambio de argumentos en la esfera pública. Si, por el contrario, desea uno averiguar de dónde obtiene su validez este paradigma de racionalidad, halla como respuesta que se trata de un paradigma implícito en la cultura pública de la sociedad. Es una argumentación circular por cuyo intermedio se postula el ideal de racionalidad intersubjetiva en términos histórico-culturales de raigambre comunitarista. A semejante postulación parece verse Rawls obligado por su pretensión de dar al liberalismo una fundamentación exenta de obligaciones cognitivistas.

No debe sorprendernos por eso que afirme Rawls: «A este factor de la cooperación social lo llamo yo *lo razonable (the reasonable)*»²³. Hemos visto ya que la noción de razonabilidad se hace necesaria como alternativa a la noción de verdad moral. En el contexto sistemático de la teoría «lo razonable» aparece como una «facultad» («*power*») de las personas —lo cual es coherente, pues personas son justamente los miembros cooperantes en el modelo elemental de la sociedad. A las personas, Rawls les atribuye, como es sabido, dos facultades básicas: la capacidad de practicar un sentido de la justicia y la capacidad de poseer una concepción del bien. A estas dos facultades les corresponden, en términos más abstractos, otras dos: al sentido de la justicia, la razonabilidad (*the reasonable*), y a la capacidad de poseer una concepción del bien, la racionalidad (*the rational*)²⁴. El término «racional» designa la actitud teleológica de los agentes individuales —también la actitud de las partes en la posición originaria—; se ajusta, pues, a un modelo de racionalidad estratégica. El término «razonable», en cambio, encierra todas las propiedades de la razón cooperativa que vimos hace un momento, y se ajusta en tal sentido a un modelo de racionalidad intersubjetiva. La razonabilidad cooperativa es el marco explicativo y normativo en el que se inscribe la racionalidad estratégica de los agentes individuales.

Con lo dicho hasta aquí tenemos ya los elementos de juicio para entender lo que significa el *consenso presupuesto*. La teoría cree disponer, como hemos visto, de un sustrato consensual y generador de consenso, pero que ella habría descubierto siendo aún sólo una «concepción autónoma» («*a free standing view*»). Supuestamente, este sustrato consensual no tendría nada que ver con el «consenso abarcador» («*overlapping consensus*») que la teoría trataría recién ahora, en una segunda fase, de generar en comunicación con las doctrinas morales comprensivas. Veamos, pues, más detenidamente en qué consiste el «consenso abarcador» y en qué medida éste logra proporcionar estabilidad a la teoría. Y veamos asimismo hasta qué punto las correcciones hechas al método de fundamentación predeterminan ya en cierto modo los rasgos del consenso que se pretende alcanzar.

²³ «El constructivismo kantiano en la teoría moral», *op. cit.*, p. 148.

²⁴ Cfr. *Political Liberalism*, II: 1.3.

II. *El consenso perseguido*

Al igual que la noción de «estabilidad», la noción de «consenso abarcador» se presta a una serie de malentendidos. Vimos al inicio, por ejemplo, que ella ha sido interpretada (malinterpretada) como un intento de negociar la validez de la teoría. Pero éste no es el único malentendido. Los hay de tan diversos tipos, que Rawls se ve obligado en su última obra a hacer una presentación negativa de esta noción, es decir, a definir el «consenso abarcador» mediante una refutación de las objeciones principales que se han formulado en su contra. De ellas nos ocuparemos en seguida.

Pero, antes de discutir las, conviene que reparemos en lo que Rawls tiene exactamente en mente al abordar esta cuestión o, dicho de otra manera, en cuál es específicamente la pregunta que se plantea. Porque, lejos de imaginarse una situación simétrica entre todas las concepciones en conflicto, lo que él se pregunta es «si la concepción política (es decir, su actual teoría de la justicia) puede convertirse en el foco de un consenso abarcador»²⁵. Este consenso tiene, por así decir, un ofertante, la teoría liberal, y un destinatario, las doctrinas comprensivas, y la forma misma de la pregunta no admite el intercambio de estos roles. Lo que Rawls se propone no es, pues, hacer que la teoría liberal entable un diálogo con las demás teorías en igualdad de condiciones, sino hacer una presentación tan persuasiva del sustrato político de su teoría que todas las demás se vean motivadas (u obligadas) a adherirse a ella. Sobre esto no parece tener dudas ni vacilaciones. Leemos así en *Political Liberalism* que la teoría liberal debe «obtener el respaldo de un consenso abarcador, es decir, de un consenso a través del cual ella llegue a ser aceptada por todas las doctrinas morales, religiosas y filosóficas rivales que se han ido desarrollando durante generaciones en un Estado más o menos justo y democrático, cuyo criterio de justicia es precisamente esta concepción política»²⁶. Las vacilaciones emergen tan sólo cuando se trata de explicar cómo las concepciones concernidas pueden aceptar la oferta del liberalismo sin tener para ello que renunciar a sus respectivas cosmovisiones. Pero esta cuestión había sido planteada ya en la primera fase de constitución de la teoría y, por lo que hemos visto, en ella se ofrecen los elementos de juicio necesarios para darle respuesta. De ser esto así, el problema del «consenso abarcador» sería un asunto propio de la definición misma del liberalismo, y no se entendería muy bien por qué Rawls insiste en considerarlo un asunto exclusivo de la segunda fase.

Pero vayamos a la caracterización del consenso que, como dijimos, tiene la forma de una refutación de objeciones. Las objeciones se resumen en cuatro: 1) que la propuesta de un consenso equivale a una mera negociación o, como

²⁵ *Ibidem*, p. 141.

²⁶ J. Rawls, «The Idea of an Overlapping Consensus», *Oxford Journal of Legal Studies*, 7 (1987), pp. 1-2.

lo expresa también Rawls, que con ella se aspira sólo a un *modus vivendi* con las demás teorías; 2) que la teoría demuestra así ser indiferente o escéptica ante la cuestión de la verdad; 3) que el liberalismo no puede dejar de ser él mismo una doctrina comprensiva, y 4) que el consenso perseguido es en realidad una utopía²⁷.

Por lo expuesto hasta aquí, no debería sernos difícil comprender por qué la primera objeción encierra un malentendido. La idea de un compromiso negociado desvirtúa la posición originaria de las partes involucradas y debilita la obligatoriedad de su asentimiento y del acuerdo logrado. Pero la aclaración de este malentendido nos muestra de manera indirecta cuán alto es el nivel de compromiso moral que Rawls asocia al asentimiento de todos. «En primer lugar —escribe—, el objeto del acuerdo, la concepción política de la justicia, es él mismo una concepción moral. Y, en segundo lugar, esta concepción es también asumida por motivos morales...»²⁸. Cuál es el asunto sobre el que ha de producirse un acuerdo —la teoría liberal— es algo que no está, pues, en discusión. Y las razones que animan a los individuos para dar su asentimiento, sean éstas del tipo que sean, pueden considerarse razones morales al menos en el sentido en que atañen a los fundamentos normativos de la vida pública.

Con respecto a la segunda objeción —que su teoría muestre indiferencia o escepticismo frente a la cuestión de la verdad—, cabe preguntarse qué tiene ella que ver con el problema del consenso abarcador. Por supuesto, la objeción es pertinente, pero es obvio que ella se refiere a la llamada primera fase de constitución de la teoría. Y el hecho de que Rawls la entienda como una objeción dirigida en contra del consenso no hace sino reforzar la sospecha de que la división en fases de la teoría es artificial. El problema que Rawls tiene aquí en mente es un problema de coherencia interna de su teoría, a saber: en qué medida es posible afirmar que la concepción política del liberalismo sea una concepción «razonable» o incluso «la más razonable»²⁹, sin formular al mismo tiempo una pretensión de verdad. Sabemos ya cuál es la finalidad de esta estrategia metodológica; se quiere así asegurar simultáneamente la consensualidad (la razonabilidad) de la teoría de la justicia y la autonomía de las concepciones éticas (su disputa por la verdad). Pero Rawls mismo parece percatarse de que también la «razonabilidad» presupone algún tipo de criterio de demarcación. Esto es lo que se pone de manifiesto justamente en el caso del consenso abarcador.

En efecto, no todas las doctrinas posibles ni tampoco todas las existentes son aquí consideradas «razonables» (*reasonable*), sino tan sólo aquellas que admiten implícitamente la vigencia de las condiciones intersubjetivas y argumentales del discurso público. «Consideraremos razonables —afirma Rawls—

²⁷ Cfr. *Political Liberalism*, IV: 3-7.

²⁸ *Ibidem*, p. 147.

²⁹ *Ibidem*, p. 150.

a las doctrinas comprensivas y a quienes las sostienen, siempre y cuando reconozcan los límites de la razón (*burdens of reason*) y las consecuencias que de allí se derivan sobre su uso público»³⁰. Pero si sólo se tiene en cuenta a las doctrinas que comparten ya *per definitionem* el sustrato de la teoría, uno se pregunta qué es lo que el consenso perseguido puede aún aportar, pues, como vemos, los destinatarios están siendo caracterizados con los mismos rasgos básicos que los ofertantes. Además, se aprecia aquí con claridad que la teoría hace uso de un criterio de demarcación muy exigente, cuya legitimidad sólo puede derivarse del carácter vinculante que se otorga al postulado de la cooperación social.

La tercera objeción sostiene que la teoría liberal no puede dejar de ser ella misma una doctrina comprensiva y universal, en la medida en que cree disponer de una base conceptual que sería común a todas las demás concepciones. Esta objeción tiene que ver con el núcleo de la propuesta consensual de la teoría, de modo que al comentarla podremos explicitar mejor algunos de sus supuestos. Recurramos, para ello, a un ejemplo que da Rawls con frecuencia, que él llama un «caso modelo (*model case*) de consenso abarcador»³¹. El modelo nos presenta a tres concepciones distintas que tratan de llegar a un acuerdo sobre el orden político que debe regirlos: una concepción religiosa (cristiana), una concepción filosófica (utilitarista o kantiana) y una concepción así llamada «pluralista». En sentido estricto, la agenda de la discusión no está abierta, pues aquello sobre lo que estas concepciones discuten es si aceptan o no la propuesta consensual de la teoría liberal, que hace las veces de ofertante. Pues bien, en el modelo (que, naturalmente, es ideal), la concepción religiosa acepta la propuesta liberal por razones que ella extrae de su propio credo confesional. La concepción filosófica, por su parte, acepta dicha propuesta por razones que se derivan de la doctrina moral que ella sostiene. Y la concepción «pluralista» la acepta por razones que son en realidad immanentes a la teoría liberal misma. Es decir, esta última posición es la del liberalismo, pero que ahora ya no aparece sólo como ofertante, sino simultáneamente como destinatario. Recién aquí el liberalismo es presentado desempeñando el mismo rol que todas las demás concepciones y, al igual que ellas, él también debe tomar posición con respecto a sí mismo. Que al hacerlo acepte la propuesta consensual que él mismo formula, no debería sorprendernos. Pero el hecho de que juegue este doble rol es un claro indicador de que, ante el «consenso abarcador», el liberalismo no sólo propone el tema de la discusión, sino proporciona además las motivaciones que son suficientes para los destinatarios. Al competir con las doctrinas comprensivas en el nivel de las visiones del mundo, el liberalismo desvaloriza las motivaciones que ellas asumen y las hace superfluas.

³⁰ *Ibidem*, p. 59.

³¹ *Ibidem*, pp. 145 y ss.

El doble rol del liberalismo se pone aún más claramente de manifiesto en la refutación de la cuarta objeción contra el «consenso abarcador», a saber, que éste sería sólo una utopía. La objeción, Rawls la interpreta como si se le estuviese exigiendo a la teoría liberal indicar en virtud de qué «fuerzas políticas, sociales y psicológicas»³² puede ella garantizar la realización del «consenso abarcador» y su estabilización en el largo plazo. Y por eso responde a la objeción haciendo un experimento mental, diseñando un proceso histórico imaginario en el que la difusión progresiva de los principios del liberalismo aparezca como un fenómeno plausible. Así como el nacimiento de la teoría liberal es justificado, retrospectivamente, recurriendo a la idea de un sistema imparcial de cooperación que estaría vigente en la tradición, así también la estabilidad del consenso es ahora postulada, prospectivamente, imaginando cómo ha de crearse y fortalecerse una tradición de «lealtad» hacia dicha idea fundacional. «Creemos así —escribe Rawls— que apenas comiencen los ciudadanos a apreciar lo que ofrece una concepción liberal, desarrollarán hacia ella una lealtad que se irá fortaleciendo a través del tiempo»³³. Vemos, pues, que el ideal del «consenso abarcador», de un consenso que va ganando siempre mayor estabilidad, no es otro que el ideal de la implantación definitiva del modelo liberal de la sociedad y la justicia. Y este ideal es asociado, retrospectiva y prospectivamente, al desarrollo de una tradición de lealtad hacia él. Se expresa allí la confianza moderna en el progreso incontenible de la libertad racional, y se presume, por consiguiente, que tarde o temprano las doctrinas comprensivas habrán de comprender que no hay modelo más «razonable» y habrán de adherirse a él.

Para terminar, quisiera resumir brevemente la tesis de este ensayo. Retomo, para ello, el hilo de mi argumentación, aunque esta vez comenzando por el final, a fin de comprobar si, como sugería al inicio, la cuestión del «consenso abarcador» tras el cual andamos es efectivamente dependiente de la cuestión del consenso presupuesto tras el cual nos hallamos. Según Rawls, el problema sistemático principal que afronta la reformulación de su teoría es el problema de la estabilidad. La solución de este problema es el «consenso abarcador». Ambos, el problema y su solución, sólo deben ser tematizados, sin embargo, en la segunda fase de constitución de la teoría, es decir, sólo después de que la teoría ha sido desarrollada en la primera fase como concepción autónoma. Pero, como acabamos de ver, el «consenso abarcador» no debe entenderse como un acuerdo entre ofertantes con iguales derechos, que pudiesen discutir, por ejemplo, sobre el objeto o las reglas de juego de su concertación. El único ofertante es aquí la teoría política liberal. El único asunto en debate es el sustrato público de la vida política, tal como éste es expresado en el ideal de un sistema imparcial de cooperación. Quienes piensan de otro modo desem-

³² *Ibidem*, pp. 158-159.

³³ *Ibidem*, p. 160.

peñan aquí el rol de destinatarios, a quienes se invita a restringir sus discrepancias a la esfera privada. Y de entre ellos, además, sólo se admite a la discusión a los que reconocen y hacen suyas las reglas intersubjetivas del uso de la razón pública. Todo esto no significa, por cierto, que el sustrato fundacional mencionado no tenga justificación, ni que el consenso sea imposible. Pero significa, sí, que la idea de un «consenso abarcador» sólo puede hacerse plausible con la ayuda de una serie de presuposiciones, todas las cuales deben ser fundamentadas en la primera fase de constitución de la teoría. Y estas presuposiciones operan implícitamente no sólo en la definición del objeto de la concertación, sino igualmente en la caracterización de los participantes en ella y en la de los roles que les corresponde desempeñar. No llega a entenderse, pues, del todo por qué se persigue una división tan tajante entre las dos fases de la teoría.

El consenso perseguido depende entonces en última instancia del consenso presupuesto. Es decir, el giro epistemológico de la teoría reposa sobre la rigurosidad de la fundamentación que pueda ofrecerse en la primera fase de su constitución. Sólo si la teoría es coherente y persuasiva como concepción autónoma podrá aspirar a obtener un «consenso abarcador». Esto fue lo que analizamos en la primera parte de nuestro trabajo. Vimos allí que Rawls trata de acceder al sustrato fundacional de su teoría sirviéndose de un tipo original de justificación. Para evitar las dificultades epistemológicas de las estrategias cognitivistas de fundamentación, recurre al método del «equilibrio reflexivo», a fin de buscar, retrospectivamente, ideas o convicciones comunes que puedan hacer las veces de punto de partida de la construcción conceptual. Por cierto, el método parece ser prometedor, pues se propone acceder a los beneficios de validación de la teoría, sin pagar el costo de una fundamentación última. Pero, como es sabido, en toda búsqueda se corre el peligro de no encontrar sino aquello que uno puede reconocer. Y así, vemos que la idea del sistema imparcial de cooperación es caracterizada utilizando las propiedades de un modelo intersubjetivo de razón que nos es suficientemente familiar por los debates actuales acerca de la racionalidad. Lo novedoso en Rawls es, por así decir, la división del trabajo entre la definición y la validación del concepto de razón: si su definición es manifiestamente contemporánea, su validación ha de ser ahora obtenida de la tradición de la sociedad democrática. Pero como tampoco en este caso puede separarse una cosa de la otra: ¿cómo evitar el riesgo de terminar validando sólo aquello que había sido ya definido de antemano?

Rawls tiene razón, sin duda, cuando afirma que el problema central de toda teoría política o moral es el problema de la «estabilidad». Pero a lo que él se refiere con esta ambigua noción es en realidad a un viejo problema que podríamos llamar el problema de la fundamentación normativa o de la capacidad de universalización de las teorías morales, aun cuando se restrinja ahora su alcance a los límites de una tradición particular. La justificación cuasi

comunitarista que Rawls nos ofrece de los fundamentos del liberalismo político no logra ser, como hemos visto, lo suficientemente persuasiva y da lugar más bien a fundadas reservas. Pero al formular estas reservas uno ingresa quizás inadvertidamente al universo consensual de razonabilidad que la teoría propugna³⁴.

³⁴ He desarrollado más ampliamente las tesis de este ensayo, en el marco del debate entre Rawls y los comunitaristas, en M. Giusti, «Die liberalistische Suche nach einem "übergreifenden Konsens"», *Philosophische Rundschau*, 41 (1994), pp. 53-73.

THEORIA

REVISTA DE TEORIA, HISTORIA Y FUNDAMENTOS DE LA CIENCIA

FUNDADA EN 1952 - SEGUNDA EPOCA
FUNDADOR: Miguel SANCHEZ-MAZAS (†)

Revista asociada a la
Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España

Vol. 11

No 26

MAYO / MAY 1996

SUMARIO / CONTENTS

EVOCACION DE MIGUEL SANCHEZ-MAZAS / EVOCATION OF MIGUEL SANCHEZ-MAZAS

Carlos PARIS (Madrid)	5-11	<i>Pequeña semblanza de Miguel Sánchez-Mazas en el aniversario de su muerte (A portrait of Miguel Sánchez-Mazas in the first anniversary of his death)</i>
	13-25	<i>Miguel Sánchez-Mazas. 50 años de actividad científica (1946-1995) (Miguel Sánchez-Mazas. 50 years of scientific activity (1946-1995))</i>

ARTICULOS / ARTICLES

Roberto TORRETTI (Santiago, Chile)	29-43	<i>Realismo científico y ciencia real (Scientific Realism and Scientific Practice)</i>
Juan Carlos GARCIA-BERMEJO OCHOA (Madrid)	45-83	<i>Comparaciones de aproximación a la verdad y de lejanía de la falsedad (Comparing statements by their closeness to the truth and their distance from the falsehood)</i>
Jon PEREZ LARAUDOGOITIA (Vitoria)	85-89	<i>New Difficulties with 'If...Then'. The Paradox of the Businessman</i>
Julia BARRAGAN (Caracas)	91-103	<i>A Type-Logic Approach to Refinement</i>
José J. MORESO, Pablo E. NAVARRO (Barcelona)	105-124	<i>Verdad y eficacia (Truth and Effectiveness)</i>
Jesús EZQUERRO, Mauricio IZA (San Sebastián)	125-145	<i>Knowledge Representation, Reflexive Reasoning and Discourse Processing</i>
Begoña VICENTE, Pablo RODRIGUEZ (Vitoria)	147-162	<i>La modalidad en el lenguaje natural: consideraciones lógicas y pragmáticas (Modality in natural language: logical and pragmatic considerations)</i>
Manuel GARCIA SERRANO (Kassel)	163-189	<i>Yo e identidad personal (Self and Personal Identity)</i>

FE DE ERRATAS / ERRORS

IN MEMORIAM

RECENSIONES Y LIBROS RECIBIDOS / BOOK REVIEWS AND BOOKS RECEIVED

CRONICAS Y PROXIMAS REUNIONES / NOTICES AND ANNOUNCEMENTS

NOTICIAS DE LA SLMFCE / NEWS OF THE SLMFCE (Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España)

REDACCION / EDITORIAL OFFICE / REDACTION

ESPAÑA/SPAIN: CALIJ-THEORIA, Alcalde José Elosegui, 275, E 20015, Apartado 1.594, 20080, San Sebastián, España. Tel.: (34 43) 29.17.25. Fax (A. Ibarra): (34 43) 31.10.56. E-mail: theoria@sf.ehu.es

EXTRANJERO/FOREIGN COUNTRIES: Asociación Cultural España-THEORIA, Case 2.730, 1211 Genève-2, Suisse.

DISTRIBUCION / DISTRIBUTION

Para suscripciones, números atrasados y cambios de dirección: /For subscriptions, back volumes and changes of address: Servicio Editorial, Universidad del País Vasco, Apartado 1.397, E 48080 Leioa, España. Tel.: (34 4) 464.77.00 (ext. 2153). Fax: (34 4) 480.13.14.

THEORIA es una revista cuatrimestral (sale en Enero, Mayo y Septiembre).

THEORIA is a four-monthly journal (issues in January, May and September).