

# **UNIVERSIDAD IBEROAMERICANA**

Estudios con Reconocimiento de Validez Oficial por Decreto Presidencial  
Del 3 de abril de 1981



LA VERDAD  
NOS HARÁ LIBRES

UNIVERSIDAD  
IBEROAMERICANA

CIUDAD DE MÉXICO ®

## **“LA TEORÍA DE JUSTICIA DE RAWLS Y SU VALIDEZ A LA LUZ DE LA CRÍTICA DE HARSANYI”**

**TESIS**

Que para obtener el grado de

**MAESTRO EN FILOSOFÍA**

Presenta

**ROBERTO GARRIDO PATRÓN**

Director: DR. FRANCISCO CASTRO MERRIEFIELD

Lectores: DR. ÁNGEL OCTAVIO ÁLVAREZ SOLÍS

DR. CARLOS MENDIOLA MEJÍA

Ciudad de México, 2017

## AGRADECIMIENTOS

Quiero expresar mi profundo agradecimiento a las personas que contribuyeron a la realización de este trabajo, en especial al Dr. Francisco Castro Merrifield, por su orientación, supervisión y guía, y sobre todo por su motivación y el apoyo brindado a lo largo de más de dos años.

Asimismo, deseo dar las gracias al Dr. Ángel Octavio Álvarez Solís, por sus críticas constructivas y por su invitación constante a ensanchar el horizonte de mi investigación para integrar más perspectivas, de una y otra escuela de pensamiento, con el objeto de tratar de enriquecer mis planteamientos.

Al Dr. Carlos Manuel Mendiola Mejía por su tiempo, y por haberme llevado a la búsqueda de la influencia de algunos filósofos clásicos en la obra de Rawls.

Un agradecimiento muy especial para mi esposa Maru por su comprensión, paciencia y apoyo incondicional.

A todos ellos, muchas gracias.

*A mi esposa Maru,  
y a mis hijos Roberto,  
María Alejandra  
y Eugenio,  
con mucho cariño.*

## ÍNDICE

### INTRODUCCIÓN

1. Justicia distributiva	5
2. Objetivo del proyecto	8
3. John Rawls	12
4. El hombre rawlsiano	13

### CAPÍTULO 1. TEORÍA DE JUSTICIA DE RAWLS

5. Rawls y su interés en la justicia	22
6. Objetivo de la teoría de la justicia	24
7. El papel de la justicia	25
8. El objeto de la justicia	26
9. La idea principal de justicia como imparcialidad	27
10. La posición original y su justificación	28
11. El utilitarismo clásico	29
12. El intuicionismo	31
13. Los dos principios de justicia y maximin	32
14. Los bienes sociales primarios y su índice	36
15. El velo de ignorancia	39
16. La racionalidad de las partes	40
17. La presentación de opciones	43
18. Las restricciones formales del concepto de lo justo	45
19. Razonamiento que conduce a los dos principios de justicia	47
20. Expresión definitiva de los dos principios de justicia	51
21. La moralidad rawlsiana	52
22. Influencia del contractualismo de Locke y Rousseau	54

23.Tres críticas a la teoría de justicia de Rawls	59
24.La teoría rawlsiana en un caso práctico	67
25.El principio de diferencia y su aplicación	69
26.La regla maximin en el principio de diferencia	71
CAPÍTULO 2. LA CRÍTICA DE HARSANYI	
27.La incertidumbre: punto de partida de Harsanyi	72
28.Elementos de la crítica de Harsanyi	73
29.Respuesta de Rawls	75
30.Comentaristas sobre la disputa Rawls-Harsanyi	79
CAPÍTULO 3. CONCLUSIONES	
31.Lo moral y lo racional en Rawls	85
32.En la búsqueda de una racionalidad máxima	86
33.El error de Rawls	89
34.Argumentación a favor de maximin	91
35.Objeciones a la crítica de Harsanyi	94
36.Conclusión desde la óptica racional estricta	95
37.¿Cómo salvar la teoría de justicia de Rawls?	96
38.Conclusión desde una óptica ampliada	98
39.La disputa Rawls-Harsanyi en la actualidad	99
40.Comentarios finales	103
BIBLIOGRAFÍA	105

# LA TEORÍA DE JUSTICIA DE RAWLS Y SU VALIDEZ A LA LUZ DE LA CRÍTICA DE HARSANYI

## INTRODUCCIÓN

La justicia es uno de los conceptos morales y políticos más importantes en la historia de la sociedad humana. La palabra deriva del latín *iustitia*. El conocido jurista romano Ulpiano, definió la justicia en el siglo III como “la constante y perpetua voluntad de dar y conceder a cada uno su derecho”<sup>1</sup>. Por su parte, la Real Academia Española define a la misma como “una de las cuatro virtudes cardinales que inclina a dar a cada uno lo que le corresponde o pertenece”<sup>2</sup>. Es claro que, en esencia, las dos versiones expresan algo similar. Pero la filosofía busca ir más allá de la etimología o de las definiciones de diccionario para considerar la naturaleza de la justicia en cuanto que virtud moral del individuo y como una cualidad política deseable en la sociedad.

En los siguientes párrafos presentaré una breve revisión sobre la justicia distributiva, tema central de este trabajo.

### 1. JUSTICIA DISTRIBUTIVA

La justicia distributiva se ocupa de la asignación imparcial de recursos entre los diversos miembros de una comunidad. Una distribución imparcial normalmente toma en cuenta el total de bienes a distribuir, el procedimiento de distribución, y el patrón de distribución que resulta.

Dado que las sociedades tienen una cantidad limitada de riqueza y recursos, la cuestión de cómo deben ser distribuidos esos beneficios surge con frecuencia. La

---

<sup>1</sup> Javier Paricio y Alejandro Fernández Barreiro, *Historia del Derecho Romano y su Recepción Europea*. Madrid: El Faro, 2002, p.25.

<sup>2</sup> Real Academia Española, *Diccionario de la Lengua Española*. Madrid: Espasa, 2014, p 901.

respuesta más común es que los bienes deberían ser distribuidos de una manera imparcial, para que cada quien reciba una 'porción justa'. Pero esto deja abierta la cuestión de lo que constituye una 'porción justa'.

Varios principios se han propuesto para determinar de qué manera deberían distribuirse los bienes. Algunos criterios son igualdad, equidad y necesidad. Si la igualdad es considerada como el criterio último para determinar qué recibe cada quién, entonces todos los bienes serán distribuidos en cantidades iguales entre las personas. Sin embargo, dado que cada persona tendrá necesidades diferentes, el resultado no será el mismo para todos.

Otra posibilidad es proceder de acuerdo al principio de equidad, y llevar a cabo la distribución en proporción a la contribución a la comunidad de cada uno de los individuos. Así, aquellos que aporten más de un recurso recibirán más. Algunos sugieren un sistema de competencia que incluya un piso de bienestar para aquellos que no pueden competir. Este tipo de sistemas combina el principio de equidad con el de necesidad, y busca compensar a las personas por su capacidad de producir asegurando al mismo tiempo que las necesidades básicas de todos estén satisfechas.

Por último, se podrían distribuir los recursos de acuerdo a la utilidad social; es decir, buscando lo que es de mayor interés para la sociedad integral.

Los distintos tipos de distribuciones promueven diferentes metas sociales. Para que una sociedad funcione efectivamente, debe mantener a sus miembros incentivados para producir bienestar de una manera eficiente y efectiva, y para sostener el bienestar de sus miembros. Los principios de justicia surgen de estas preocupaciones. La distribución igualitaria se percibe como la que da un mayor sentido de pertenencia a todos los miembros de la sociedad. La equidad fomenta la motivación para producir, a ser premiado por la productividad personal. Finalmente, la distribución de acuerdo a la necesidad buscaría que las necesidades básicas de todos estuvieran satisfechas lo cual no sólo es bueno para el individuo, sino que también tiende a disminuir la violencia social.

Dado que estos ideales con frecuencia están en tensión unos con otros, uno de ellos es normalmente propuesto como el criterio central de distribución. Así,

dependiendo de qué principio de distribución se adopte, el sistema económico se caracterizará por igualdad, competencia, o protección social.

Algunos consideran que lo que hace justo a una distribución es el resultado final, mientras que para otros lo es las reglas que se asignan para determinar la distribución. Incluso en aquellos casos en los que el resultado logre ser una distribución imparcial de recursos, los procedimientos para llegar a esa distribución podrían ser considerados como injustos por algunos. Por el contrario, un procedimiento justo podría resultar en una distribución injusta.

En su *Teoría de la Justicia*<sup>3</sup>, John Rawls afirma que el lugar de nacimiento de la persona, su estatus social, y las influencias familiares son resultado de la suerte y por tanto no deberían influenciar en alto grado la cantidad de beneficios que recibimos en la vida. Él sostiene que el papel de la justicia distributiva es limitar la influencia de la suerte de manera que los bienes puedan ser distribuidos de una manera más justa para beneficio de todos.

Otros piensan que la justicia distributiva es un asunto de diseñar reglas que los individuos debieran seguir para la adquisición y transferencia de recursos y beneficios. Por tanto, el objetivo de la justicia distributiva no es lograr un resultado particular de patrón de distribución, sino asegurar un proceso de intercambio que sea justo. Algunos otros piensan que la justicia distributiva debería ser una cuestión tanto de proceso como de resultado. Consideran que el proceso de distribución debe ser imparcial para que las personas sientan que han recibido un resultado justo.

En algunos casos, el bien a distribuir no es un beneficio, sino una carga, como podría ser el caso de las penalidades por no obedecer la ley, o los gastos médicos de las personas que no cuentan con seguro.

Cuando las personas perciben que sus beneficios no están a la par de los que recibieron otros en un lugar similar, desarrollan un sentido de injusticia. Cuando las personas tienen la sensación de que están en desventaja con respecto a otros, pueden desear confrontar al sistema que ha dado lugar a estos desbalances de oportunidades

---

<sup>3</sup> John Rawls, *Teoría de la Justicia*, Tr. María Dolores González. México: FCE, 1978.

y de bienestar. Esto es particularmente delicado cuando hay grupos de personas que no han visto satisfechas sus necesidades más fundamentales.

La redistribución de beneficios puede algunas veces ayudar a aliviar las tensiones sociales y permitir una sociedad más estable. Sin embargo, si el proceso de redistribución es visto como ilegítimo, los problemas sociales pueden regresar.

Así, un sistema de justicia distributiva efectivo e imparcial será siempre una importante balanza para regular el sistema de cooperación y distribución de cargas y beneficios sociales de manera que los miembros de la comunidad puedan prosperar en un clima adecuado de armonía social.

## 2. OBJETIVO DEL PROYECTO

El tema a investigar en este trabajo es la teoría de justicia de John Rawls, profundizando en el principio de diferencia, y en el uso de la regla maximin<sup>4</sup>, el cual fuera duramente reprobado por el académico de la doctrina utilitarista John Harsanyi<sup>5</sup>. El propósito de mi investigación es concluir si la propuesta de Rawls se sostiene a la luz de la crítica del economista.

Por tanto, inicio en esta *Introducción* con una reflexión filosófica sobre la concepción antropológica asumida por Rawls en su teoría de justicia, sistema que denomina *justicia como imparcialidad*. Más adelante, en el Capítulo 1 expongo los elementos centrales del aparato conceptual propuesto por Rawls en su obra clásica *Teoría de la Justicia*, a la que también me referiré indistintamente como *TJ*, con especial atención a sus dos principios de justicia, y al papel que juega la regla maximin como sostén fundamental de uno de ellos: el principio de diferencia. Adicionalmente, reflexionaré sobre la función primordial del principio maximin como criterio de elección de principios permanentes de justicia. En el Capítulo 2 analizaré la crítica de Harsanyi, presentada en su publicación *¿Puede el principio maximin funcionar como una base*

---

<sup>4</sup> La regla maximin supone que la alternativa preferible es aquella en la que el resultado más bajo es el menos dañino de los ofrecidos por las opciones disponibles. Es un concepto creado en matemáticas y en teoría de juegos. Ver Garth Kemerling, *A Dictionary of Philosophical Terms and Names*, 2001. <<http://www.philosophypages.com/dy/m2.htm#maxi>> (consulta 12 ago. 2014).

<sup>5</sup> John Harsanyi recibió el premio Nobel de economía en 1994.

*para la moralidad? Una crítica a la teoría de John Rawls*<sup>6</sup>, cuya apreciación negativa argumenta que el uso de esta regla conduce a decisiones y consecuencias equivocadas. Asimismo, presentaré la óptica de otros autores contemporáneos sobre la validez de la crítica del economista. Por último, en el Capítulo 3 expondré mis reflexiones personales sobre los argumentos de Rawls y de Harsanyi en pro y en contra del uso de la regla maximin, con el objeto de concluir con mi propia argumentación sobre la validez del uso de este criterio de priorización en la teoría de justicia rawlsiana.

Rawls plantea su sistema de justicia en oposición directa a la moral utilitarista, que asume que la medida de lo correcto es aquello que genere mayor bienestar para el mayor número de personas. En efecto, el autor dice en el prólogo de su obra “esta teoría (de justicia) parece ofrecer otra explicación sistemática de la justicia que es superior, al menos así lo sostengo, al utilitarismo dominante tradicional.”<sup>7</sup>

Para plantear su sistema de justicia, Rawls establece que la estructura básica social será considerada el sujeto principal, entendiendo a ésta como la manera en que las principales instituciones sociales distribuyen derechos y deberes, y la forma en que determinan la división de las ventajas sociales entre los miembros de la población.

La organización de la obra de Rawls que nos ocupa es bastante lineal. En la primera parte del libro, el autor diseña una posición teórica que denomina la Posición Original (*PO*), y argumenta que las partes en ella preferirán su sistema de justicia como imparcialidad, por encima de la justicia basada en la moral utilitaria. Para entender el argumento desde la *PO*, es necesario comprender las motivaciones de las partes, que son especificadas cuidadosamente por Rawls. La segunda parte de *Teoría de la Justicia* verifica la consistencia entre los principios de justicia como imparcialidad y los puntos de vista de las instituciones angloamericanas de justicia. La tercera parte del texto analiza la estabilidad de una sociedad organizada bajo los principios de Rawls, en donde el filósofo argumenta que existirá una congruencia importante en esa sociedad

---

<sup>6</sup> John Harsanyi, “Can the maximin principle serve as a basis for morality? A critique of John Rawls’ theory”, *American Political Science Review*, 59, 1975, p. 594-606.

<sup>7</sup> *Op. cit.*, p.10.

entre los puntos de vista de las personas sobre lo que consideran justo y lo que valoran.

La Posición Original pregunta cuáles principios de justicia social serán elegidos por las partes representantes de los individuos, donde se tiene en cuenta que dichas partes serán totalmente ajenas —debido al velo de ignorancia— respecto a las preferencias de las personas que representan. Rawls describe el velo de ignorancia cuando escribe “nadie sabe cuál es su lugar en la sociedad, su posición, clase o status social; nadie sabe tampoco cuál es su suerte en la distribución de ventajas y capacidades naturales, su inteligencia, su fortaleza, etc.”<sup>8</sup>

El objetivo de Rawls es demostrar que las partes tendrán preferencia por los principios de justicia como imparcialidad por encima de los de la moral utilitarista. Rawls no busca convencer al lector de que su propuesta es la mejor de todas las concepciones de justicia que se han argumentado a lo largo de la historia; él busca específicamente dejar claro que su teoría es superior a la de la maximización de la utilidad. A lo largo de esta investigación buscaré concluir si el autor logra su objetivo.

En términos generales, los principios de justicia propuestos por Rawls son:

Principio 1. Cada persona debe tener un derecho a la más amplia libertad, compatible con una libertad similar para los otros.

Principio 2. Parte 1: las estructuras sociales que dan forma a la distribución de bienes deben satisfacer el requerimiento de igualdad imparcial de oportunidades. Parte 2:

Principio de diferencia: las desigualdades sociales y económicas deben resolverse teniendo como criterio el mayor beneficio de los más desaventajados. Este principio de diferencia está basado en la regla maximin, que establece que el criterio de selección de alternativas de acción deberá elegir siempre el camino en el que el menos favorecido tiene la mejor posición.

Harsanyi presenta cuatro argumentos en favor del utilitarismo y en contra de la aplicación del principio maximin favorecido por el pensador americano. Harsanyi propone en esencia que el utilitarismo es un criterio más eficaz, y que el principio

---

<sup>8</sup> *Ibidem.*, p.25

maximin lleva a conclusiones equivocadas. En el Capítulo 2 expondré en detalle los elementos de esta reprobación.

Esta crítica de Harsanyi a su vez ha generado reacciones por parte de otros filósofos de actualidad, dando así evidencia de la vigencia del debate sobre este tema. A manera de ilustración, considero revelador lo que el pensador contemporáneo belga Philippe Van Parijs comenta:

Pocos elementos de la filosofía política de John Rawls han probado ser tan definitorios de una época como lo que él de manera un tanto extraña llamó 'el principio de diferencia'. Ninguno ha ejercido tanta influencia fuera del círculo de los filósofos académicos. Y difícilmente alguno ha generado tantas confusiones o provocado tantas acaloradas controversias.<sup>9</sup>

Es, pues, evidente que el principio de diferencia, y su reproducción del criterio maximin dentro de su redacción, han sido generadores de gran polémica intelectual, incluso fuera del ámbito filosófico, dado que la crítica que nos ocupa proviene de un prominente economista. Adicionalmente, Van Parijs alude a la gran cantidad de literatura secundaria que ha generado este principio rawlsiano, y refiere<sup>10</sup> ocho obras concretas que se han dedicado a reflexionar sobre el tema. Otra verificación de la actualidad del debate es el hecho de que año tras año se continúen publicando ensayos sobre este tópico en múltiples medios especializados<sup>11</sup>, a más de cuatro décadas de distancia de la crítica original de Harsanyi.

En sinopsis, la investigación de este trabajo será revisar el contenido fundamental de la justicia como imparcialidad de Rawls, y en particular los dos principios de justicia que la definen, con énfasis en el principio de diferencia, por su uso de la regla maximin. Complementaré la revisión con la crítica de algunos autores contemporáneos sobre la

---

<sup>9</sup> Philippe Van Parijs, "Difference Principles", en *Oxford Companion to Rawls*, Ed. Samuel Freeman. New York: Cambridge University Press, 2003.

<sup>10</sup> *Ibidem.*, nota p.234.

<sup>11</sup> Para una de las más recientes publicaciones, ver: Sarkar, Sumit and Chakraborty, Soumyakanti, Does Rawls' 'Original Position' Induce Fairness? Experimental Findings on Selection Criteria in a Discrete Nash Demand Game Played from Behind the 'Veil of Ignorance' (March 18, 2016). SSRN: <http://ssrn.com/abstract=2749496> or <http://dx.doi.org/10.2139/ssrn.2749496>.

propuesta del pensador americano. Después revisaré en detalle la crítica de Harsanyi. Para entender su relevancia, también estudiaré la respuesta del propio Rawls al economista utilitarista, y expondré el punto de vista de expertos que han estudiado la disputa y dado su veredicto acerca de quién tiene la razón. Mi objetivo final será concluir con mi propia argumentación filosófica si la *Teoría de la Justicia* de Rawls se sostiene, con su uso de la regla maximin, frente a la crítica planteada por Harsanyi.

### 3. JOHN RAWLS

John Rawls nació en la ciudad de Baltimore, Estados Unidos, en 1921. Hijo de un prominente abogado, estudió su licenciatura en Harvard, y años más tarde se doctoró en Princeton en filosofía moral, en 1950. Una parte de su disertación fue la base del primer artículo que publica: *Guía de un Proceso de Decisión para Ética*<sup>12</sup>. Éste fue un primer intento para abordar la pregunta central de su teoría, que maduraría más tarde: ¿qué proceso de decisión podemos imaginar que nos ayude a resolver diferencias entre las partes de una manera justa?

Después de enseñar en Cornell y Massachusetts Institute of Technology, Rawls ingresó a Harvard como profesor de filosofía en 1962, institución en la que permaneció casi cuarenta años. A lo largo de su carrera, se concentró en dos temas: la filosofía moral y la filosofía política.

Con la publicación de *TJ* en 1971, Rawls adquirió un reconocimiento considerable en el medio académico. Sus obras adicionales más importantes son *Liberalismo Político*<sup>13</sup>, *El Derecho de Gentes*<sup>14</sup> y *La Justicia como Equidad: Una Reformulación*<sup>15</sup>. El pensador americano fallece en 2002.

---

<sup>12</sup> John Rawls, "Outline of a Decision Procedure for Ethics", en *Philosophical Review*. April 1951, 60 (2), p.177-197.

<sup>13</sup> John Rawls, *El Liberalismo Político*, Tr. Vicente Orellana Aranda. Barcelona: Crítica, 2004.

<sup>14</sup> John Rawls, , *El Liberalismo Político*, Tr. Vicente Orellana Aranda. Barcelona: Crítica, 2004..

<sup>15</sup> John Rawls, *La Justicia como Equidad: Una Reformulación*, Tr. Andrés de Francisco. Barcelona: Paidós, 2012.

#### 4. EL HOMBRE RAWLSIANO

Antes de exponer la teoría de justicia de Rawls, me parece relevante preguntarnos qué concepción del hombre tiene en mente el autor cuando diseña su esquema para la distribución justa de los bienes. Para tener un punto de referencia que me permita analizar este tema, tomaré como base las dos concepciones clásicas, y opuestas entre sí, de la naturaleza humana: la de Hobbes y la de Rousseau.

En lo que respecta a Hobbes, en su tesis expuesta en *Leviatan*<sup>16</sup>, en 1651, el pensador considera que el hombre es un ser constantemente motivado por apetitos y aversiones, con ciertas metas en mente que siempre intentará alcanzar. Dado que los hombres frecuentemente desearán el mismo bien, como alimento o un lugar donde vivir, estarán en conflicto permanente unos con otros. Si estos apetitos fueran finitos, el conflicto no sería muy problemático; sin embargo, los deseos humanos no tienen saciedad y por lo tanto la lucha de poder será continua.

Al respecto, Hobbes se refiere de la siguiente manera:

De esta igualdad de capacidades surge la igualdad en la esperanza de alcanzar nuestros fines. Y, por lo tanto, si dos hombres cualesquiera desean la misma cosa que, sin embargo, no pueden ambos gozar, devienen enemigos; y en su camino hacia su fin (que es principalmente su propia conservación, y a veces sólo su delectación) se esfuerzan mutuamente en destruirse o subyugarse. Y viene así a ocurrir que, allí donde un invasor no tiene otra cosa que temer que el simple poder de otro hombre, si alguien planta, siembra, construye, o posee asiento adecuado, puede esperarse de otros que vengan probablemente preparados con fuerzas unidas para desposeerle y privarle no sólo del fruto de su trabajo, sino también de su vida, o libertad. Y el invasor a su vez se encuentra en el mismo peligro frente a un tercero.<sup>17</sup>

---

<sup>16</sup> Thomas Hobbes, *Leviatán*, Tr. Manuel Sánchez Sarto. México: FCE, 2002.

<sup>17</sup> *Ibidem*, p. 52.

Pienso que esta visión refiere al hombre como su peor enemigo, y del que más cuidado deberá tener, pues inevitablemente habrá más de un individuo deseando la misma cosa; por tanto, el conflicto violento será inevitable, e incluso posiblemente mortal. Es por ello, dice Hobbes, que esta igualdad en el estado natural hace necesaria la presencia de un poder común para mediar entre los hombres y distribuir los recursos; de otra manera no habría más que una guerra constante en la que “la vida del hombre es solitaria, pobre, desagradable, brutal y breve”<sup>18</sup>. El autor introduce aquí de manera contundente la necesidad de establecer una autoridad que controle esta tendencia intrínseca del ser humano hacia la mutua destrucción.

Por su parte, Rousseau tiene una concepción del ser humano completamente opuesta. El filósofo francés postula en 1755, en su *Discurso Sobre el Origen de la Desigualdad*<sup>19</sup>, que el hombre en su estado natural es fundamentalmente impulsado por dos motivaciones principales: la misericordia y la auto-preservación. En su estado natural, el hombre existe sin la conciencia del bien o del mal, y es feliz en esencia. Sólo lo afecta su necesidad consciente de alcanzar la perfección, la cual obliga a las personas a cambiar en el tiempo.

Para Rousseau la única desigualdad entre los seres humanos es la fuerza física. Sin embargo, nos dice el filósofo, en las sociedades modernas la creación de leyes ha corrompido al hombre natural y ha generado nuevas formas de desigualdad que no son propias de su estado de origen.

La posición de Rousseau sobre la bondad natural del hombre se ejemplifica en esta cita:

¿Mas qué es la generosidad, la clemencia, la humanidad, sino compasión dirigida al más débil, al culpable, o al ser humano en general? Incluso la benevolencia y la amistad son, si las juzgamos correctamente, solamente los efectos de la compasión, constantemente fijada con un

---

<sup>18</sup> *Ibidem*, p. 104.

<sup>19</sup> Jean-Jacques Rousseau, *Discurso sobre el Origen y los Fundamentos de la Desigualdad entre los Hombres*, Tr. Ángel Pumarega. Valencia: Tilde, 1997.

objetivo particular: ¿cómo pueden ser diferente el desear a otra persona que no sufra dolor ni ansiedad, de el desearle felicidad?”<sup>20</sup>.

El texto alude manifiestamente a la compasión como virtud fundamental del ser humano; es el ver por el otro como si fuera yo mismo.

Paso ahora a estudiar la concepción antropológica en la teoría de justicia de Rawls. Para ello, me apoyaré en la metodología hermenéutica, acudiendo a la obra respectiva para buscar evidencias de la presencia de las nociones del hombre de Hobbes y de Rousseau en la teoría rawlsiana. Primeramente, acudiré al texto *Teoría de la Justicia* en busca de las citas ofrecidas por el autor sobre los dos filósofos de la antropología y analizaré el contexto en que son citados para intentar concluir si cada una de esas menciones implica o no una evidencia. Como segunda fuente, revisaré los supuestos expresados sobre las características del hombre por el propio Rawls en el texto, que pudieran revelar una concepción antropológica implícita; es decir una concepción que a mi parecer pueda ser interpretada de manera concluyente como más cercana a uno de los dos prototipos de referencia. Mi hipótesis es que a través de estos dos métodos resultará claro si el hombre rawlsiano tiende más hacia al modelo bondadoso de Rousseau, o hacia la visión de Hobbes asociada al egoísmo.

Comienzo con Hobbes, a quien Rawls cita a lo largo de su obra en cinco ocasiones. La primera cita dice: “Pese a su grandiosidad (...) el *Leviatán* de Hobbes, hace surgir algunos problemas especiales”<sup>21</sup>. Rawls no abunda en los ‘problemas especiales’, pero especulo que el comentario es indicativo de que la teoría política completa de Hobbes tiene algunas discrepancias significativas con la de Rawls<sup>22</sup>. No profundizaré más en el tema, dado que este tópico está fuera del alcance de la actual investigación. En todo caso, lo relevante es que este comentario específico de Rawls no apoya la visión antropológica de Hobbes. Por lo tanto, es descartado como evidencia.

La siguiente cita pertenece a la sección 38, *El Imperio de la Ley*:

---

<sup>20</sup> *Ibidem.*, p. 52

<sup>21</sup> *Op. cit.*, nota p. 24.

<sup>22</sup> Ver Joseph Grcic, *Rawls and Hobbes on Political Power*, <<https://periodicos.ufsc.br/index.php/ethic/article/viewFile/17444/16066>>.

Los acuerdos para cumplir la ley de manera voluntaria son inestables. Por esta razón se supone que un soberano coercitivo es siempre necesario, aunque en una sociedad bien ordenada las sanciones no sean severas y acaso no necesiten imponerse. No obstante, la existencia de una maquinaria penal eficaz sirve para garantizar la seguridad de las personas. Esta proposición y el razonamiento que implica podemos considerarlos como tesis de Hobbes<sup>23</sup>.

Aquí Rawls reconoce de manera manifiesta la necesidad fundamental de la existencia del soberano coercitivo para el adecuado funcionamiento de su sistema de justicia. El texto incluso contiene una nota a pie de página que hace referencia al *Leviatán*, indicando capítulos específicos. Ésta es, pues, una primera evidencia clara sobre la presencia de la concepción hobbesiana en la obra que ocupa esta investigación.

Hobbes es mencionado por tercera ocasión en la sección 42, *Algunas Consideraciones Acerca de sistemas Económicos*, que trata del problema del aislamiento, entendido como aquella situación en la que:

...el resultado de las decisiones de muchos individuos, tomadas aisladamente, es peor para cada uno que otro curso de acción, aunque dando por sentada la conducta ajena, la decisión de cada persona fuera perfectamente racional. Es el caso del dilema del prisionero, cuyo ejemplo clásico es el estado de naturaleza de Hobbes<sup>24</sup>.

Veo aquí otra referencia al estado natural de Hobbes, ahora en el contexto de circunstancias en las que las elecciones individuales e independientes de cada hombre llevan a soluciones sub-óptimas para la colectividad. Rawls considera que el aislamiento es una fuente potencial de desestabilización de su propuesta de justicia. Argumento que esta alusión constituye una segunda evidencia de la presencia de la concepción egocentrista antropológica.

---

<sup>23</sup> *Op. cit.*, p. 227.

<sup>24</sup> *Ibidem*, p.252

La cuarta mención a Hobbes sucede en la sección 52, *Argumentos en Pro del Principio de Imparcialidad*:

La obligación de cumplir una promesa es una consecuencia del principio de imparcialidad. El papel de las promesas es análogo al que Hobbes atribuía al soberano. Del mismo modo que el soberano mantiene y estabiliza el sistema de cooperación social manteniendo públicamente una serie de penas, así los hombres en ausencia de acuerdos coercitivos establecen y estabilizan sus empresas privadas dando su palabra unos a otros.<sup>25</sup>

Nuevamente destaca la referencia al soberano, ahora para respaldar la obligación de cumplir las promesas que los hombres formalizan; la premisa es que sin la existencia de un sistema coercitivo, el hombre tratará de evadir los compromisos que le signifiquen un costo mayor al que aceptaría de buena voluntad. Encuentro aquí una tercera evidencia del modelo hobbesiano.

La quinta y última cita se da en la sección 76, *El Problema de la Estabilidad Relativa*, la cual señala que: “Podemos considerar al soberano hobbesiano como un mecanismo agregado a un sistema de cooperación que sin él sería inestable”<sup>26</sup>.

Una vez más, se hace presente la figura de la máxima autoridad para brindar permanencia a la teoría rawlsiana. Ésta constituye, a mi parecer, la cuarta evidencia de los rasgos hobbesianos.

En sinopsis, he encontrado cuatro evidencias referidas a las menciones del propio Rawls que apuntan hacia una concepción egoísta del hombre. La primera, relacionada con la necesidad de un sistema penal que garantice la seguridad de las personas. La segunda, ligada al requerimiento de resolver el problema del aislamiento y su efecto desestabilizador. En tercer término, la referencia al requerimiento de un soberano coercitivo que asegure el cumplimiento de las promesas. Finalmente, la llamada al soberano para brindar estabilidad al sistema de justicia.

Avanzo ahora a las citas sobre Rousseau en *Teoría de la Justicia*, quien es mencionado por Rawls en diez ocasiones. De ellas, cuatro las considero relacionadas a

---

<sup>25</sup> *Ibidem*, p.317

<sup>26</sup> *Ibidem*, p.449

la tradición del contrato social, por lo que se excluirán del tema en cuestión; otra más es sobre la tolerancia religiosa; dos sobre la influencia de Rousseau en la doctrina moral de Kant; y una sobre la teoría freudiana que refiere al interés de las personas en la búsqueda de la justicia como una forma reactiva. Dado que estas ocho menciones están relacionadas con temas diferentes al que nos ocupa, las descartaré como posibles evidencias. Por tanto, me enfocaré en la revisión de las dos menciones restantes, para examinar si constituyen evidencias sobre la concepción del hombre bondadoso.

La primera es en la sección 69, *El Concepto de una Sociedad Bien Ordenada*. Aquí Rawls hace referencia a las dos tradiciones del aprendizaje moral: la teoría de Freud y la del pensamiento racional ilustrada por Rousseau. La conclusión del texto es que “esta tradición considera los sentimientos morales como consecuencia natural de una plena apreciación de nuestra naturaleza social”<sup>27</sup>. Interpreto que, si bien este comentario es claramente alusivo al hombre bueno, Rawls no toma una posición al respecto, sino que simplemente lo menciona. Por lo tanto, no me parece adecuado considerarla una evidencia.

La siguiente cita es en la sección 70, *La Moral de la Autoridad*. Aquí el texto refiere al *Emilio*<sup>28</sup> de Rousseau, al decir que:

Podemos suponer que los padres quieren al niño y que, con el tiempo, el niño llega a querer a sus padres y a confiar en ellos. ¿Cómo se produce este cambio en el niño? Para contestar a esta pregunta, doy por sentado el siguiente principio psicológico: el niño llega a querer a sus padres sólo si antes ellos le quieren manifiestamente a él.<sup>29</sup>

Aunque aquí Rawls toma una postura de apoyo al escritor francés, al dar por sentado el principio que toma prestado de *Emilio*, interpreto esta manifestación del amor referido del niño a sus padres más como un reflejo de los rasgos egoístas del ser humano que de las características bondadosas con las que asociamos el modelo

---

<sup>27</sup> *Ibidem*, p. 419

<sup>28</sup> Jean-Jacques Rousseau, *Emilio, o de la Educación*. Tr. Mauro Armíño. Madrid: Alianza, 2005

<sup>29</sup> *Op. cit.*, p.419

antropológico de Rousseau, pues el esquema es que el niño brindará amor a los padres, sólo si él primero lo recibe de ellos. Por tanto, argumento que no debe ser considerado como una evidencia en favor de la presencia de la antropología del filósofo francés.

En sinopsis, la revisión de las dos citas referidas a Rousseau, me hace concluir que ninguna de ellas las puedo interpretar como manifestaciones del hombre rousseauiano en la postura de Rawls.

He concluido esta primera exploración basada en las citas formales. Me parece evidente que el hombre de Hobbes tiene un peso específico concreto en la doctrina rawlsiana, constituido por las cuatro evidencias ya mencionadas. Por el contrario, insisto en que el modelo de Rousseau no tiene una sola mención de peso como concepción antropológica en el texto que pueda ser considerada como una verificación.

Paso ahora a la segunda búsqueda hermenéutica. Aquí intentaré identificar algún rasgo importante del aparato conceptual de Rawls que tenga implicaciones claramente antropológicas, o una parte de su texto que exprese una concepción concreta de la naturaleza humana, aun cuando no esté conectada a ninguno de los autores que he propuesto como referencia. Aquí mi objetivo es buscar en la narrativa construida por Rawls, partes del texto que puedan ser interpretadas como manifestaciones del concepto antropológico que el propio autor expresa, aun cuando fuese de manera indirecta.

Una primera evidencia surge al comienzo de *TJ* en la sección 1, *El Papel de la Justicia*, cuando Rawls dice: “La desconfianza y el resentimiento corroen los vínculos del civismo, y la sospecha y la hostilidad tientan al hombre a actuar en formas que de otro modo evitaría”<sup>30</sup>. Ésta es una clara alusión al hombre egoísta y preso de sus pasiones, que tiende a rebasar los límites cívicos cuando percibe que el otro no le ha dado el trato que se merece. Me parece relevante que el filósofo resalte estos rasgos negativos generales del ser humano de una manera tan clara en la misma primera sección en la que describe el papel de la justicia, con lo que sugiere que un sistema de justicia no puede ser concebido de manera aislada sin considerar las características del

---

<sup>30</sup> *Ibidem*, p. 20.

sujeto humano y sin tomar las medidas necesarias, al establecer el aparato, para que este lado pasional de los individuos no desvirtúe los objetivos trazados por el autor.

Previo al siguiente punto, deseo enfatizar que, a mi juicio, la columna vertebral de la presente obra de Rawls es precisamente la construcción sistemática de la argumentación que sostiene la elección de los dos principios de justicia, por encima de otras teorías de justicia que el mismo autor presenta, que culmina en la sección 26 de la obra. Esta argumentación será analizada en detalle más adelante. Dicho lo anterior, la segunda referencia antropológica ocurre precisamente en la parte crucial de la obra, la sección 26, *El Razonamiento que conduce a los Dos Principios de Justicia*. Aquí Rawls escribe:

“Existe una relación entre los dos principios y la regla maximin para escoger en condiciones de incertidumbre. Esto es evidente a partir del hecho de que los dos principios son aquellos que escogería una persona al proyectar una sociedad en la cual su enemigo hubiera de asignarle su lugar”<sup>31</sup>.

Pienso que esta referencia alude de manera contundente al comportamiento humano hobbesiano en dos aspectos: primero, asume que en el curso de las relaciones humanas se generan conflictos, y que se puede llegar al grado de crear enemigos que desean hacer daño al otro; segundo, me resulta revelador de la concepción rawlsiana del hombre el que se presente el razonamiento para la elección de los dos principios de justicia - la parte culminante del aparato teórico de justicia-, asumiendo de manera íntimamente ligada la posibilidad de que el hombre no solamente esté en manos de su enemigo como representante social, sino que además este último cuente con la capacidad hipotética unilateral de asignarle su lugar en la sociedad.

Resumo ahora los hallazgos hermenéuticos de las dos búsquedas. En la primera, que consideré las citas del autor de *Teoría de la Justicia* sobre Hobbes y Rousseau, ya expuse que las referencias al filósofo inglés son de un peso claro y fuerte. El peso es otorgado por las cuatro evidencias encontradas, que refieren a la necesidad de un

---

<sup>31</sup> *Ibidem*, p.150.

sistema penal que garantice la seguridad de las personas, al requerimiento de resolver el problema del aislamiento y su efecto desestabilizador, y a la obligatoria presencia de un 'soberano coercitivo' para asegurar el cumplimiento de las promesas, y brindar estabilidad al sistema de justicia. Por el contrario, las citas sobre Rousseau se concentran en su labor relacionada con la teoría del contrato social, y no hay ninguna referencia específica de peso sobre su concepción antropológica. En cuanto a la segunda búsqueda de evidencias dentro del texto de la argumentación de Rawls para sostener su aparato conceptual, se encontraron dos claras alusiones a las características hobbesianas del hombre: una que resume en una breve frase las consecuencias de la desconfianza y el resentimiento, y otra que resalta la consideración del enemigo en el razonamiento que conduce a la elección de los dos principios de justicia. Por lo tanto, a partir de la hipótesis que presenté inicialmente, considero que el propio texto da evidencias suficientes para concluir con certeza que la concepción del hombre rawlsiano es mucho más cercana a la de Hobbes que a la de Rousseau.

Existe otra cualidad interesante para caracterizar la antropología de Rawls: el eje de la dimensión histórico-vacío, en el que los dos extremos son, por un lado, el hombre situado en una época específica, en el contexto de una sociedad concreta que vive condiciones particulares; y, por el otro, el individuo como agente moral abstracto, que no se encuentra ligado a ningunos fines determinados, ni a una comunidad particular, ni a ningún momento histórico, que sería lo que denominaré antropología *vacía*. Para investigar este eje, acudo al texto de Jesús Rodríguez, *El Igualitarismo Liberal de John Rawls*.<sup>32</sup> Rodríguez es claro cuando dice que Rawls "ha partido de una noción de la persona no definida en términos nacionales, comunitarios o territoriales", aludiendo a la falta de contexto histórico de las personas colocadas en la posición original de su esquema. Así, concluyo que la antropología rawlsiana es completamente vacía.

En sinopsis, pienso que la antropología de Rawls es una combinación de la noción hobbesiana egoísta del hombre, el cual está al mismo tiempo desposeído de

---

<sup>32</sup> Jesús Rodríguez, *El Igualitarismo Liberal de John Rawls*. México: UAM, 2010.

comunidad, territorio específico, o contexto histórico concreto. En otras palabras, la antropología rawlsiana es egoísta y vacía.

## **CAPITULO 1. TEORÍA DE JUSTICIA DE RAWLS**

En este apartado expongo el objetivo que se propone Rawls al desarrollar de su teoría de justicia, los conceptos que elaboró para poder construir su propuesta, y el recorrido intelectual que siguió para llegar a la formulación y eventual elección de los dos principios de justicia enunciados en su aparato conceptual. Para ello me respaldaré en su obra *Teoría de la Justicia*.

### **5. RAWLS Y SU INTERÉS EN LA JUSTICIA**

Los extraordinarios logros de Rawls como académico, autor y profesor pueden atribuirse a diversos factores, como su gran capacidad intelectual, y virtudes tales como una muy desarrollada aptitud de pensamiento sistemático, sobresaliente memoria, y una sana actitud crítica frente a su propia obra que le generaba insatisfacción productiva.

Rawls concentró sus energías y esfuerzos en dos profundas preguntas de gran importancia personal para él: la posibilidad de que un orden social sea justo, y que la vida humana, con sus condicionantes fácticos, valga la pena vivirse.<sup>33</sup> Su enorme deseo de encontrar la respuesta a estas preguntas, evidente en sus escritos, mantuvo concentrado a Rawls durante su larga vida de trabajo.

Rawls recibe una invitación de Harvard para ingresar como profesor en 1961. Recordemos que, políticamente, la década de los años sesenta estuvo dominada por la guerra de Vietnam. Desde su origen en 1955, Rawls estimaba que esta guerra era injusta, y defendió en público su punto de vista reiteradamente. Incluso participó en una conferencia anti-guerra en Washington, en mayo de 1967. En 1969 enseñó el curso *Los problemas de la guerra*, en el cual hacía críticas a la guerra de Vietnam.

---

<sup>33</sup> Delfín Grueso (compilador), *Rawls, Legado de un Pensamiento*. Cali: Ed. Universidad del Valle: 2005, p.54.

En el fondo, Rawls estaba profundamente consternado por comprender qué fallas en la sociedad americana podían explicar la continuación tan violenta de una guerra que él consideraba claramente injusta, y de qué manera los ciudadanos podrían mostrar su oposición a este conflicto. Con respecto a la primera pregunta, concluye que el problema principal es el hecho de que la riqueza está distribuida de manera muy dispar, obedeciendo al modelo utilitarista, y que ésta se utiliza fácilmente como influencia política para favorecer con legislación parcial a los privilegiados del sistema, perpetuando el ciclo de la desigualdad. En específico, el proceso político de Estados Unidos está estructurado de manera que ofrece a los individuos y a las corporaciones adineradas, incluyendo las de la industria de la defensa, la oportunidad de tener dominio político a través de sus contribuciones a los partidos relevantes.

Escrita durante esos años, *Teoría de la Justicia* refleja esos pensamientos:

Aquellos dotados y motivados de manera similar deberían tener básicamente la misma oportunidad de alcanzar las posiciones de autoridad política sin tener en cuenta su clase económica y social....Históricamente uno de los principales defectos del gobierno constitucional ha sido la falla en asegurar el justo valor de las libertades políticas...Las disparidades en la propiedad y la riqueza que exceden con creces lo que es compatible con la igualdad política han sido generalmente bien toleradas por el sistema legal.<sup>34</sup>

Esta crítica es ampliada en un texto posterior *The Basic Liberties and their Priority*.<sup>35</sup> Con respecto a la segunda pregunta, Rawls la afrontó de manera más activa y visible. En aquel momento muchos jóvenes no estaban dispuestos a hacer su servicio militar, que era obligatorio hasta los 26 años, e ir a la guerra. El gobierno americano había decidido no reclutar estudiantes con buen promedio académico, otorgando con ello un poder inusual a los profesores. ¿Por qué algunos estudiantes deberían ser tratados mejor que otros, sobre todo considerando que los padres adinerados tienen ventaja clara para asegurar un lugar para sus hijos en una universidad de su elección? Si los hombres jóvenes deben ser obligados a ir a la guerra, entonces al menos los hijos de

---

<sup>34</sup> *Op. cit.*, p.214

<sup>35</sup> John Rawls & Sterling M. McMurrin, *Liberty, Equality and Law: Selected Tanner Lectures on Moral Philosophy*. Salt Lake City: University of Utah Press, 1987.

los privilegiados deberían compartir este mismo camino. Si no es necesario que acudan a la guerra todos los jóvenes, entonces el número requerido debería ser determinado por la suerte.

Junto con varios de sus colegas en el Departamento de Filosofía de Harvard, Rawls defendió esta postura y propuso su adopción. Finalmente, la moción se sometió a votación y fue derrotada, pero el debate en relación con la guerra de Vietnam continuó en Harvard muchos años. Me parece que las huellas de estos pensamientos están presentes en *Teoría de la Justicia* cuando el texto refiere:

No merecemos el lugar que tenemos en la distribución de dones naturales, como tampoco nuestra posición inicial en la sociedad. Igualmente problemático es el que merezcamos el carácter superior que nos permite hacer el esfuerzo por cultivar nuestras capacidades, ya que el carácter depende, en buena parte, de condiciones familiares y sociales afortunadas en la niñez, por las cuales nadie puede atribuirse mérito alguno.<sup>36</sup>

Así, Rawls echa por tierra la teoría de la meritocracia, según la cual cada quien recibe los frutos de su esfuerzo y trabajo, dentro de un marco justo de oportunidades equitativas. Su visión de que la distribución de talentos y posición inicial es una lotería será uno de los fundamentos para proponer el principio de diferencia, que busca compensar a los menos afortunados en ese juego de azar.

Después de estos antecedentes que permiten comprender algunas de las vivencias que influenciaron a Rawls en su interés y concentración en el tema de la justicia, paso a revisar su propuesta.

## **6. OBJETIVO DE LA TEORÍA DE LA JUSTICIA**

Desde el prefacio, Rawls es claro al indicar que su intención es postular una propuesta que sea superior a la concepción de justicia utilitarista, hasta entonces reinante. Refiere el texto:

---

<sup>36</sup> *Op. cit.*, p.106

Lo que he tratado de hacer es generalizar y llevar la teoría tradicional del contrato social representada por Locke, Rousseau y Kant, a un nivel más elevado de abstracción...Más aún, esta teoría parece ofrecer otra explicación sistemática de la justicia que es superior, al menos así lo sostengo, al utilitarismo dominante tradicional.<sup>37</sup>

A lo largo de la obra el autor desarrolla un minucioso procedimiento para establecer las condiciones que debe cumplir un sistema de justicia, y confronta sus dos principios contra otras doce teorías tradicionales –incluyendo desde luego al utilitarismo en sus diversas formas- para mostrar que sus dos principios serán elegidos por las partes bajo el marco dado.

## **7. EL PAPEL DE LA JUSTICIA**

Para Rawls “la justicia es la primera virtud de las instituciones sociales, como la verdad lo es de los sistemas de pensamiento.”<sup>38</sup> Cada persona tiene una inviolabilidad fundamentada en la justicia que ni siquiera la búsqueda utilitarista del bienestar social integral puede golpear. Refiriendo a la primacía de la libertad sobre el bienestar económico, “la justicia niega que la pérdida de libertad para algunos se vuelva justa por el hecho de que un mayor bien es compartido por otros.”<sup>39</sup>

Por lo tanto, en un sistema social justo, las libertades correspondientes a la igualdad como ciudadano quedan firmemente establecidas. Los derechos fundamentados en la justicia no dependen de negociaciones políticas ni de intercambio de intereses sociales.

En una sociedad, encontraremos típicamente tanto conflicto de intereses como identidad de intereses. Existe la identidad de intereses en tanto que la cooperación social hace factible para la agrupación de personas una vida mejor que la que cada quien lograría con esfuerzos individuales aislados. Y se da el conflicto de intereses puesto que cada individuo tiene sus criterios personales de valoración, y no serán indiferentes respecto a cómo han de compartirse los beneficios de la cooperación

---

<sup>37</sup> *Ibidem*, p.10

<sup>38</sup> *Ibidem*, p.17

<sup>39</sup> *Idem*.

social. Los principios de justicia, el resultado final de la propuesta rawlsiana, son de gran peso, pues su objetivo es “proporcionar un modo para asignar derechos y deberes en las instituciones básicas de la sociedad y definen la distribución apropiada de los beneficios y las cargas de la cooperación social.”<sup>40</sup>

El concepto de sociedad bien ordenada es un presupuesto clave en la propuesta de Rawls, quien dice que “una sociedad está bien ordenada no sólo cuando fue organizada para promover el bien de sus miembros, sino cuando también está eficazmente regulada por una concepción pública de justicia.”<sup>41</sup> Esto tiene dos condiciones: cada individuo acepta y sabe que las demás personas aceptan los mismos principios de justicia, y además las instituciones sociales básicas aplican de manera satisfactoria estos principios, y es del conocimiento público que lo hacen. Así, el interés general por la justicia genera vínculos de amistad cívica y limita la persecución de otros fines. Podemos decir que la concepción pública de justicia viene a ser el atributo esencial de una sociedad bien ordenada.

Respecto a las instituciones sociales, éstas serán justas cuando no hagan distinciones parciales entre los individuos al asignarles derechos y obligaciones básicas, y cuando las normas determinen un adecuado equilibrio social entre proyectos individuales que compitan con las ventajas de la cooperación social.

La doctrina utilitarista se limita a plantear una concepción de justicia meramente distributiva de los beneficios económicos. Rawls propone que no es posible evaluar una teoría de justicia solamente por su papel distributivo. Debemos tomar en cuenta sus implicaciones más amplias, ya que una concepción de justicia es preferible a otra cuando sus resultados generales son más beneficiosos y deseables para los miembros de la sociedad.

## **8. EL OBJETO DE LA JUSTICIA**

Es bien sabido que se dice de diversas cosas que son justas o injustas: no solamente la ley, o las instituciones sociales, sino también las acciones de las personas,

---

<sup>40</sup> *Ibidem.*, p.18

<sup>41</sup> *Idem.*

incluyendo las decisiones. También se refiere a la justicia como virtud ética cuando se habla de la disposición de los individuos para resolver de la manera más equitativa posible los conflictos que la vida les presenta. Para Rawls, el objeto primario de la justicia es lo que él denomina “estructura básica social”, entendida como “el modo en que las grandes instituciones sociales distribuyen los derechos y deberes fundamentales y determinan la división de las ventajas de la cooperación social.”<sup>42</sup> Más específicamente, en el contexto de una democracia moderna, el término refiere a la constitución política y a los principales ordenamientos económicos y sociales.

Son las grandes instituciones sociales contemporáneas las que definen los derechos y obligaciones de los individuos, y repercuten sobre sus expectativas de vida, favoreciendo unas posiciones iniciales sobre otras. Estas son desigualdades particularmente profundas. No solamente determinan fácticamente las oportunidades del hombre, sino que no pueden ser justificadas apelando a conceptos de mérito de ningún tipo. Por tanto, insiste Rawls, es a las inequidades de la estructura básica social, tal vez inevitables, a las que deben aplicar en primer lugar los principios de justicia.

El ámbito de la teoría de justicia rawlsiana queda delimitado por dos fronteras. La primera es que no considera la justicia de las prácticas específicas de las instituciones sociales concretas, pues éstas son una abstracción que permite al autor desarrollar su modelo. Tampoco examina la justicia de los acuerdos contractuales particulares. La segunda limitación es muy significativa: solamente desarrolla los principios de justicia que regularían a una sociedad bien ordenada. Esta última premisa define a la propuesta como una teoría ideal, y Rawls lo reconoce cuando dice: “empezaré con la teoría ideal porque creo que ofrece la única base para una comprensión sistemática de los problemas más apremiantes.”<sup>43</sup>. En otras palabras, el autor decide hacer unas simplificaciones que le permitan llegar a establecer los principios adecuados de justicia, en aras de que, a partir de ellos, aun cuando hayan sido generados desde una realidad

---

<sup>42</sup> *Ibidem.*, p.20

<sup>43</sup> *Ibidem.*, p.22

idealizada, se puedan establecer normas que sean aplicables a contextos reales de la compleja sociedad democrática.

## **9. LA IDEA PRINCIPAL DE JUSTICIA COMO IMPARCIALIDAD**

La noción del contrato social, o contrato original, de Locke, Rousseau y Kant, es introducida por Rawls, aclarando que no debemos pensar en el contrato social como el mecanismo requerido para ingresar a una sociedad en particular. La propuesta es que los principios de justicia para la estructura básica social sean el objeto del acuerdo original, y que dichos principios normen todos los acuerdos posteriores y las formas de gobierno que puedan establecerse. A este modo de operar de la regulación social, Rawls lo denomina *justicia como imparcialidad*, y con ése término nos referiremos a su propuesta integral.

En este planteamiento teórico un grupo determinado de individuos tiene el deber de decidir, de una vez y para siempre, lo que para ellos significará justo o injusto en la sociedad. Así, la elección que los hombres racionales hagan en esta condición hipotética define y formaliza los principios de justicia.

En la justicia como imparcialidad, la posición original de igualdad refiere al estado de naturaleza inicial de la teoría del contrato social. A este momento fundante Rawls le denomina la *posición original*, uno de los conceptos más relevantes de su andamiaje conceptual.

La posición original no está concebida como una condición históricamente real. Es un recurso puramente hipotético que conduce a una determinada concepción de justicia. La posición original está caracterizada por ciertos condicionantes. El más importante es el *velo de ignorancia*, en virtud del cual las partes representantes en la situación original eligen los principios con desconocimiento total sobre tres temas vitales: nadie sabe cuál es su lugar en la sociedad, su situación o nivel social; nadie sabe cuál es su fortuna en la distribución de ventajas y capacidades naturales, inteligencia y fortaleza; y, por último, nadie está enterado de su propia concepción del bien, ni de sus tendencias psicológicas particulares. Así, dadas estas condiciones de la posición original y la simetría de las relaciones entre las partes, la situación inicial resulta equitativa entre los individuos, en tanto que seres morales. Esto identifica lo

apropiado del nombre *justicia como imparcialidad*, pues expresa que los principios de justicia se eligen y acuerdan en una condición inicial justa.

Para pasar de la concepción hipotética de la posición original a la realidad legal de la sociedad, se asume que después de haber elegido los principios de justicia, las partes determinarán una constitución y un poder legislativo que aplique las leyes en concordancia con los principios de justicia acordados.

Una premisa esencial es que las partes representantes en la posición original son racionales y mutuamente desinteresados. La interpretación de racionalidad debe ser en el sentido económico tradicional, de acuerdo con el cual se hace uso de los medios más eficaces para intentar alcanzar los fines dados.

## **10. LA POSICIÓN ORIGINAL Y SU JUSTIFICACIÓN**

Una concepción de justicia será más aceptable que otra, si personas razonables puestas en la posición original escogieran sus principios en lugar de los de la segunda, para llevar a cabo el papel de la justicia. Por tanto, debemos averiguar qué principios sería racional adoptar desde la situación originaria. Así, el objetivo del proceso de elección de los principios de justicia será descartar aquéllos principios sobre los que supiéramos cosas improcedentes desde la perspectiva de la justicia, aun cuando fuese en principio racional postularlos como candidatos a ser elegidos.

Una manera de verificar la razonabilidad de los principios elegidos es ver si corresponden a las convicciones que tenemos de la justicia. En específico, si el aplicar estos principios nos condujera a hacer los mismos juicios que ahora tenemos de manera intuitiva sobre la estructura básica social, tendremos confianza en nuestra elección, pues esos principios tendrán la capacidad de ser consistentes con nuestras más firmes convicciones.

La tesis es que este proceso iterativo para definir la condición inicial, a partir de condiciones débiles que se van reforzando, nos llevará al punto en que la propia concepción de posición inicial será capaz de generar los principios de justicia que sean consistentes con ella. Esta solución se encontrará en el estado de cosas llamado “equilibrio reflexivo”. Es equilibrio porque nuestros principios y nuestros juicios coinciden; y también es reflexivo dado que sabemos a qué principios se ajustan

nuestros juicios reflexivos. Es de notar que este equilibrio no es necesariamente estable todo el tiempo; podría ser modificado por los resultados de un examen adicional o por los casos específicos que pudieran invitarnos a modificar nuestros juicios.

## 11. EL UTILITARISMO CLÁSICO

Dado que la teoría de justicia rawlsiana se desarrolla en oposición al utilitarismo, es prudente examinar brevemente esta doctrina. Existen varias formas de utilitarismo; sin embargo, Rawls refiere a la versión clásica como se entiende en la obra de Sidgwick<sup>44</sup>. La idea principal de esta tesis es que cuando las instituciones más relevantes de la sociedad están organizadas de tal modo que obtienen el mayor equilibrio neto de satisfacción distribuido entre todos los miembros de la sociedad, entonces la sociedad está correctamente ordenada y es, por tanto, justa.

Esta noción sobre la sociedad puede parecer atractiva a primera vista, y es posible suponer que la concepción de justicia más racional es la utilitaria. Para comprender esto, pensemos en el paradigma generalmente aceptado de que cada hombre actúa de manera correcta, al menos cuando otros no son afectados, cuando trata de obtener el mayor beneficio posible mientras persigue sus fines racionales. Si esta visión es adecuada para un individuo, ¿por qué no habría de serlo para una asociación de hombres? Bajo esta lógica, el bienestar de la sociedad se construiría a partir de la satisfacción de los sistemas de deseos de los múltiples individuos que la conforman. Así, mediante esta reflexión, se llega de manera natural al principio de utilidad: una sociedad está correctamente organizada cuando sus instituciones maximizan el resultado neto de satisfacción.

Una consideración adicional hace todavía más atractiva la propuesta utilitarista. Las dos nociones principales de la ética son lo bueno y lo justo. La configuración de una teoría ética será entonces definida por el modo de concebir y de conectar estas dos ideas fundamentales. Parecería que el camino más sencillo de relacionarlas es el que proponen las teorías teleológicas como el utilitarismo: el bien es definido de manera independiente de lo justo, y lo justo es definido como aquello que maximiza el bien. Por

---

<sup>44</sup> Henry Sidgwick, *The Methods of Ethics*, New York: Dover, 1966

tanto, serán justas aquellas instituciones y actos que produzcan el mayor bien. Así, esta tesis tiene un atractivo natural pues parece evidente que las cosas debieran ordenarse de modo que conduzcan al mayor bien posible.

Sin embargo, esta propuesta tan tentadora tiene algunas fracturas fundamentales. El rasgo más sorprendente de la tesis utilitarista es que no es relevante, excepto de manera indirecta, la forma en que se distribuya la suma de satisfacciones entre los individuos; tampoco importa cómo un individuo distribuye sus satisfacciones en el tiempo. La distribución correcta en los dos casos es aquella que produzca la máxima satisfacción. Es más, en sí misma, ninguna distribución es mejor que otra, excepto en el caso extremo teórico de que una distribución más igualitaria sea preferida para romper vínculos.

Así, la forma más natural de llegar al utilitarismo es postular para la sociedad en su conjunto el principio de elección racional del individuo. Esto resulta problemático en tanto que asume que un observador imparcial perfectamente racional se identifica con los otros miembros de la sociedad y tiene la experiencia de sus deseos como si éstos fueran propios, y les asigna un valor en la estructura única de deseos, cuya satisfacción intentará maximizar el dirigente legislador a través del ajuste de las reglas de interacción y distribución social. Esta idea de cooperación social funde a todos los individuos de la sociedad en uno, a través de la imaginación del espectador imparcial. En consecuencia, el utilitarismo no respeta la individualidad de las personas y las considera de manera simplista como una gota en el océano de la utilidad agregada.

## **12. EL INTUICIONISMO**

Rawls dedica unas páginas a describir el intuicionismo, dado que esta filosofía moral es considerada por algunos como una teoría de justicia. El intuicionismo sostiene que, una vez alcanzado cierto nivel de generalidad, ya no se pueden determinar criterios constructivos de orden superior que cumplan con el objetivo de determinar el acento adecuado en otros principios alternos de justicia. De hecho, las teorías intuicionistas tienen dos rasgos particulares: el primero es que están constituidas por diversos primeros principios que pueden en ocasiones estar en conflicto, llegando a dar soluciones contrapuestas en ciertos casos; el segundo es que no proponen un método

específico ni reglas de prioridad para valorar entre sí los diversos principios generales. Así, el intuicionismo propone que hagamos una valoración intuitiva de los principios disponibles, para llegar a través de un ejercicio de aproximación a la decisión más correcta en cada caso. Cuando llega a haber reglas de prioridad, éstas son consideradas como secundarias, y en general no proporcionan una ayuda de peso para lograr un juicio definitivo.

Existen algunas propuestas ligadas al intuicionismo. Por ejemplo, la que sostiene que los conceptos de lo justo y de lo bueno son simplemente imposibles de analizar. Otras indican que los principios morales, cuando llegan a ser expresados propiamente, aportan realmente poco, puesto que expresan proposiciones evidentes acerca de supuestos morales que de hecho son en sí legítimos.

En opinión de Rawls existen muchas clases de intuicionismo, las cuales pueden ser identificadas de acuerdo al nivel de generalidad de sus principios. Por ejemplo, el intuicionismo de sentido común propone un grupo concreto de normas, las cuales son aplicadas a cada caso particular de justicia, con énfasis específico en algún tema: salarios justos, o el pago de impuestos, y así sucesivamente.

Un problema general de las teorías de justicia intuicionistas es que las personas que emiten su juicio tenderán a apoyar los criterios que promuevan sus intereses particulares. Adicionalmente, las ideas de las personas sobre la justicia se verán fuertemente influenciadas por las costumbres culturales de su entorno y las expectativas sociales del momento, con el reto implícito de la dificultad de generar criterios para juzgar la justicia de la costumbre y la legitimidad de las expectativas socialmente construidas.

El intuicionismo postula que la complejidad de los hechos morales es superior a nuestra capacidad de dar una explicación integral de nuestros juicios y que, por tanto, requiere de un numeroso repertorio de principios alternos. Para Rawls, el único camino para refutar con autoridad al ambiguo intuicionismo es proponer criterios éticos específicos que reconozcan y expresen los valores que a nuestro parecer son correctos. También sería una refutación adecuada proponer los tipos de principios éticos constructivos que, de acuerdo con el intuicionismo, no existen; un ejemplo de dichos principios son los que postula la justicia como imparcialidad de Rawls.

### 13. LOS DOS PRINCIPIOS DE JUSTICIA Y MAXIMIN

La primera sección de la teoría de la justicia rawlsiana se divide en dos partes medulares: 1) una interpretación de la posición original y una estructuración de los distintos principios disponibles en ella para su elección, y 2) un razonamiento que permita concluir cuál de estos principios será elegido.

Rawls enuncia, de manera provisional, los dos principios de justicia sobre los que piensa que habrá consenso en la posición original. A medida que avanza en la solidificación de sus argumentos, llega a la formulación definitiva. Los dos principios provisionales son:

*Primero:* Cada persona ha de tener un derecho igual al esquema más extenso de libertades básicas que sea compatible con un esquema semejante de libertades para los demás.

*Segundo:* Las desigualdades sociales y económicas habrán de ser conformadas de modo tal que a la vez que: a) se espere razonablemente que sean ventajosas para todos; b) se vinculen a empleos y cargos asequibles para todos.<sup>45</sup>

Hay dos frases que parecen ambiguas en el segundo principio: 'ventajosas para todos' y 'asequibles para todos'. Más adelante se dará la expresión final de los dos principios que buscará hacerlos más claros y específicos.

El primer principio se encarga de definir y asegurar las libertades básicas iguales, mientras que el segundo concierne a las desigualdades económicas y sociales. Las libertades básicas comprendidas por el primer principios son: la libertad política (el derecho a votar y a ser elegible para ocupar puestos públicos) y la libertad de expresión y de reunión; la libertad de conciencia y de pensamiento; la libertad de la persona que incluye la libertad frente a la opresión psicológica, la agresión física y el desmembramiento (integridad de la persona); el derecho a la propiedad personal y la libertad respecto al arresto y detención arbitrarios, tal y como está definida por el

---

<sup>45</sup> *Op. cit.*, p.67-68

concepto de estado de derecho. Estas libertades deberán ser iguales para todos los ciudadanos.

El segundo principio se acciona, en esta primera aproximación, con respecto a la distribución del ingreso y la riqueza y a establecer organizaciones que hagan uso de las diferencias de autoridad y responsabilidad. Lo que postula es que la distribución del ingreso y la riqueza no necesariamente debe ser igualitaria; sin embargo, debe de ser ventajosa para todos, y al mismo tiempo los puestos de autoridad y poder deben ser accesibles a todos.

Estos principios deberán ser dispuestos en un orden serial, estableciendo la prioridad del primer principio sobre el segundo. Esta ordenación significa algo vital para Rawls: las violaciones a las libertades básicas iguales preservadas por el primer principio no pueden ser justificadas ni resarcidas mediante mayores ventajas sociales y económicas. Estas libertades tienen un papel central, y solamente pueden ser objeto de límites y negociaciones cuando entren en pugna con otras libertades básicas. Por tanto, ninguna de estas libertades es absoluta.

Los dos principios son un caso especial de una concepción más amplia de justicia que se expresa en lo siguiente:

Todos los valores sociales –libertad y oportunidad, ingreso y riqueza, así como las bases del respeto a sí mismo- habrán de ser distribuidos igualitariamente a menos que una distribución desigual de alguno o de todos estos valores redunde en una ventaja para todos.<sup>46</sup>

La injusticia consistirá, entonces, en las desigualdades que no benefician a todos. Esta concepción requiere ser interpretada para su adecuada explicación. Primeramente, supongamos que la estructura básica social en encarga de distribuir determinados bienes primarios, es decir, cosas que se asume que todo individuo racional desea. Asumamos que los principales bienes primarios a disposición de los miembros de la sociedad son derechos, libertades, oportunidades, ingreso y riqueza.

Imaginemos una situación hipotética inicial en la cual todos los bienes primarios sean divididos de manera igualitaria: cada persona tiene derechos y deberes

---

<sup>46</sup> *Op. cit.*, p.69

semejantes, y el ingreso y la riqueza se comparten en montos iguales. Esta situación de referencia permite a Rawls establecer un criterio para emitir un juicio sobre las mejoras: si ciertas desigualdades de riqueza y diferencias de poder resultaran en mejoras para todos en esa situación inicial teórica, entonces habría acuerdo general sobre la concepción de justicia.

Por su parte, el hecho de que los dos principios se apliquen a las instituciones de la estructura básica social tiene ciertas implicaciones. En específico, los derechos y libertades básicos a los que se refieren estos principios, son aquellos que están definidos por las normas públicas de la estructura básica. La libertad de las personas está determinada por los derechos y obligaciones determinados por las principales instituciones de la sociedad. El primer principio exige que ciertos tipos de reglas, aquellas que definen las libertades básicas, sean aplicados a cada persona de manera equitativa. En cuanto al segundo principio, el hecho de que se aplique a formas institucionales implica que refiere a las expectativas de individuos representativos; es decir, ninguno de los principios se aplica a la distribución de bienes particulares a individuos particulares que puedan ser referidos con sus nombres concretos. El caso de un criterio para asignar ciertos bienes a personas necesitadas conocidas, no se encuentra dentro del espacio de los principios, pues éstos están diseñados para regular acuerdos institucionales fundamentales.

Es de notar que el segundo principio enfatiza la exigencia de que cada individuo mejore su condición en el caso de que haya desigualdades permisibles dentro de la estructura básica. Esto implica que para cada individuo representativo tiene que ser razonable el preferir sus perspectivas con la desigualdad, al lugar que tendría sin ella. De igual manera, no está permitido justificar las diferencias de ingresos o en cargos de autoridad y poder con el argumento de que las desventajas de quienes están en una posición se compensan por las ventajas de quienes están en otra. Y menos aún se pueden comprometer de esta manera las restricciones a la libertad.

Respecto a maximin, deseo enfatizar que este criterio es el que Rawls propone para ser utilizado por las partes representantes al momento de elegir los principios de justicia de entre las alternativas que les serán presentadas en la posición original. En otras palabras, si las partes aceptan este criterio, se preguntarán '¿cuál de estos

diferentes principios alternativos de justicia es la 'solución maximin', que brinda una concepción mínima de justicia que cumple con las condiciones estrictas que Rawls establece para los principios. En la sección 18, *El Razonamiento que conduce a los Dos Principios*, expondré la manera en que Rawls articula su argumentación a favor de la aplicación de maximin, dadas las especificaciones establecidas para la posición original.

Maximin no es, obviamente, el único criterio de elección plausible. La maximización de la utilidad es otra regla de elección alterna, y su uso conduce directamente a la preferencia por el sistema utilitarista.

Me parece importante distinguir este uso de la regla maximin, de su otra presencia, dentro del principio de diferencia, el cual establece que las desigualdades en bienestar económico serán aceptables siempre y cuando se brinde prioridad a la maximización de los beneficios para los menos favorecidos. Es importante notar que la crítica de Harsanyi va dirigida al uso de maximin como regla de elección de principios de justicia, y no a su aplicación en el principio de diferencia.

#### **14. LOS BIENES SOCIALES PRIMARIOS Y SU ÍNDICE**

Rawls aborda el tema de los bienes sociales primarios a partir de la presentación de algunas dificultades del utilitarismo. Cuando el principio de la utilidad es puesto en práctica en la estructura básica, éste nos exige la maximización de la suma algebraica de las utilidades esperadas para todas las posiciones relevantes. En este sentido, es claro que el utilitarismo asume que existe una medición bastante exacta de estas expectativas. Y no solamente se debe tener una medida numérica para cada representante, sino que estas medidas deben hacer sentido en las comparaciones interpersonales. Pero el solo hecho de que pretendamos realizar este tipo de comparaciones interpersonales de niveles de bienestar, no implica que entendamos las bases de estas comparaciones o que debamos aceptarlas como correctas. Para resolver este tema necesitamos dar cuenta de estos juicios y dar a conocer los criterios que en ellos subyacen.

En temas de justicia social deberíamos intentar encontrar algunos fundamentos objetivos para estas comparaciones, bases sólidas que las personas puedan reconocer y aceptar. Rawls postula que el utilitarismo actual no es capaz de dar una respuesta

adecuada a estas dificultades. Es por ello que el principio de utilidad hace unas exigencias a nuestra capacidad de cálculo del equilibrio de las ventajas, que se queda corto en la definición de un tribunal concreto de apelaciones para asuntos de justicia.

Sin embargo, las verdaderas dificultades del utilitarismo, dice Rawls, se encuentran en otra parte. Esto es porque, aun cuando se pudiesen realizar comparaciones interpersonales satisfactorias, estas comparaciones deberían expresar valores que tenga sentido investigar. En efecto, resulta claramente ambiguo preferir un bien sobre otro simplemente porque se le pueda calcular con mayor precisión. Otro problema mayor para el pensador es, en primer lugar, si es la felicidad total, o un promedio de ella, lo que ha de maximizarse.

En contraste, el principio de diferencia busca establecer de dos maneras diferentes unas bases objetivas para efectuar las comparaciones interpersonales. Primeramente, mientras podamos identificar al representante menos aventajado, a partir de ese momento solamente serán necesarios juicios ordinales del bienestar. La ventaja sobre el utilitarismo es que sabremos desde qué posición deberá juzgarse el sistema social: la del menos favorecido. No es relevante en qué medida este representante de referencia esté peor situado que los demás; lo importante es que ya no serán necesarias más comparaciones interpersonales. Así, el principio de diferencia exige menos esfuerzos y cálculos en nuestros juicios de bienestar. Es más, no tendremos nunca que calcular una suma de ventajas que deban resultar en una cantidad ordinal.

Hay otra ventaja que ofrece el principio de diferencia en la forma de una simplificación para la base de las comparaciones interpersonales. Estas comparaciones se hacen en función de las expectativas de bienes sociales primarios, las cuales son definidas como el índice de estos bienes que un individuo representante pueda esperar. Así, las expectativas de una persona serán superiores a las de otra si este índice es mayor.

Respecto de los bienes primarios, estos son las cosas que un hombre racional desearía tener, como son derechos, libertades, oportunidades y poderes, además de ingresos y riquezas. En particular, las libertades y las oportunidades serán definidas por las reglas de las principales instituciones, y la distribución del ingreso y la riqueza estará regulada por ellas.

Se asume que aun cuando los planes racionales de las diferentes personas contemplen fines diferentes, todas ellas requerirán para su ejecución algunos tipos de bienes primarios naturales y sociales. Los planes de cada persona serán diferentes, dado que las capacidades individuales difieren, al igual que las circunstancias personales y familiares de cada quien. Así, las expectativas de los individuos representantes serán definidas por el índice de bienes primarios que tendrán a su disposición. Situadas en la posición original, las personas sabrán que sí prefieren tener más, y no menos, de estos bienes.

Claramente una de las dificultades más importantes de este planteamiento es la formación del índice de bienes primarios, pues no es obvio cómo deberán sopesarse sus diferentes componentes. Sin embargo, si asumimos que los principios están serialmente ordenados, el problema se simplifica considerablemente. Por una parte, las libertades y las oportunidades son siempre iguales para todos, por lo que estas variables no necesitan ser comparadas con otras. Por otra, los bienes primarios que sí varían en su distribución son los poderes, la dimensión de la autoridad, el ingreso y la riqueza. Dado que el punto de referencia relevante es el nivel más bajo de bienestar, el único problema de cálculo es el del índice del grupo menos aventajado. No es necesario definir ninguna posición más favorecida, mientras estemos seguros de que en realidad es más aventajada. De hecho, la correlación observada en la realidad entre poderes y riqueza nos simplifica el cálculo, por lo que el problema del índice se reduce realmente a valorar sólo el caso de los menos aventajados. El procedimiento sugerido por Rawls es ponernos “en el lugar del individuo que representa al grupo (de los menos favorecidos), preguntándonos qué combinación de bienes sociales primarios sería racional que prefiriese.”<sup>47</sup> El autor admite que este proceso se apoya inevitablemente en apreciaciones intuitivas, aunque el objetivo del ejercicio es remplazar los juicios morales por los de la prudencia racional enfocándonos en la intuición.

Se podría objetar que las expectativas no debieran medirse a través de un índice de bienes primarios, sino como las satisfacciones que se esperan obtener cuando se pongan en marcha los planes que hagan uso de estos bienes. Sin embargo, justicia

---

<sup>47</sup> *Op. cit.*, p.97

como imparcialidad toma una perspectiva diferente. No le compete el uso que las personas hagan de los derechos y las oportunidades de que disponen para medir y mucho menos maximizar las satisfacciones que logran. En vez de eso, supone que los agentes son personas racionales capaces de ajustar su concepción del bien a su situación individual. Así, razona Rawls, las expectativas de bienestar se definen como el índice de bienes primarios que una persona puede razonablemente esperar.

Pienso que la propuesta del índice de bienes primarios llega a cierta claridad dentro del ámbito abstracto de este concepto, especificando que el único caso importante a resolver es el de los menos favorecidos, lo cual acepto. Sin embargo, me parece que el autor no es capaz de ofrecer pautas para el camino que pudiera llevar un cálculo concreto, aun cuando fuese solamente un estimado. Este concepto del índice de bienes primarios también ha sido sujeto de críticas.

## **15. EL VELO DE IGNORANCIA**

La intención de la posición original es instaurar un procedimiento equitativo que permita asegurar que los principios acordados, cualesquiera que sean éstos, resulten justos. Para lograr esto, asumimos que las partes representantes están situadas bajo un velo de ignorancia, cuyo efecto es que no saben cómo las diversas alternativas de distribución de bienes primarios afectarán sus situaciones particulares, de tal manera que se ven obligadas a sopesar los principios exclusivamente en base a consideraciones generales.

Ante todo, la condición fundamental, fuertemente restrictiva, del velo de ignorancia sobre las partes es que

nadie conoce su lugar en la sociedad, su posición o clase social; tampoco sabe cuál será su suerte en la distribución de talentos y capacidades naturales, su inteligencia y su fuerza, etc. Igualmente, nadie conoce su propia concepción del bien, ni los detalles de su plan racional de vida, ni siquiera los rasgos particulares de su propia psicología, tales como su aversión al riesgo o su tendencia al pesimismo o al optimismo. Más todavía, supongo que las partes no conocen las circunstancias particulares de su propia

sociedad. Esto es, no conocen su situación política o económica, ni el nivel de cultura y civilización que han sido capaces de alcanzar.<sup>48</sup>

El único hecho específico que conocen las partes, es que su sociedad está regida por los principios de justicia, con lo que esto implica. También conocen los rasgos generales acerca del funcionamiento de la sociedad humana, los elementos de la política y los principios de la economía. Asimismo, conocen y comprenden los fundamentos de la organización social y las leyes de la psicología humana. En otras palabras, conocen todos los hechos generales que inciden en la elección de los principios de justicia.

Una característica importante deseable de la concepción de justicia es que debería generar su propio soporte. Es decir, sus principios deberían ser tales que, una vez integrados a la estructura básica social, los miembros de la sociedad tendieran a desarrollar un sentido de la justicia y el deseo de actuar de acuerdo a sus principios. Esto permitiría que el esquema de justicia fuera estable. Esta premisa es aceptable en la posición original.

La posición original no debe ser imaginada como una asamblea histórica que incluyera en un momento dado a aquéllos que vivieran en un tiempo específico, puesto que no es una reunión de personas reales y posibles. Si concebimos la posición original de esta manera, la concepción ya no sería una guía natural de la intuición y estaría vacía de un sentido claro. La posición original, como referencia conceptual, tiene que ser interpretada de forma que en cualquier momento podamos adoptar su perspectiva, y llegar a la elección de los mismos principios, independientemente de quien adopte esta perspectiva. Así, el velo de ignorancia constituye una condición esencial para la consistencia en la adopción de los principios de justicia más adecuados.

Las razones a favor del velo de ignorancia son diversas. Se desea definir la posición original de forma que lleguemos a la solución deseada. Si se permite a las partes tener conocimiento de su situación particular, la elección estará prejuiciada por las contingencias injustificadas. Si la posición original ha de producir acuerdos justos, las

---

<sup>48</sup> *Ibidem*, p.136-137.

partes deberán estar equitativamente colocadas, y ser tratadas equitativamente en tanto que personas morales. Uno de los principales objetivos del velo de ignorancia es corregir la arbitrariedad del mundo mediante el ajuste de las circunstancias de la situación contractual originaria. Es más, si al procedimiento de elegir los principios le exigimos la condición de unanimidad de los representantes, ésta será posible gracias a la exclusión del conocimiento en las partes, arribando así a una concepción de justicia que representa una genuina reconciliación de intereses. Deseo enfatizar que Rawls propone la unanimidad, como una condición de la aprobación de los principios, lo que manifiesta una cualidad de igualdad en el trato de los miembros de la sociedad, pues cada uno de ellos tendrá el mismo peso en la decisión, a través de sus representantes.

## **16. LA RACIONALIDAD DE LAS PARTES**

En todo momento Rawls ha supuesto que las partes son racionales. Esto significa que, al elegir principios, cada quien hace todo lo posible por promover sus intereses personales. Pero simultáneamente –algo muy importante- la premisa ha sido que las partes no conocen su concepción del bien. Esto quiere decir que, aun sabiendo que tienen algún plan racional de vida, no conocen los intereses y fines particulares que desean promover. La pregunta que surge es ¿cómo pueden, entonces, elegir la concepción de justicia que les será más favorable a sus intereses racionales particulares? Rawls postula que aceptan una explicación del bien que refiere a una clara preferencia a tener más bienes sociales primarios que a tener menos, un supuesto fundamental en la estructura conceptual rawlsiana. Esta suposición es racional en sí, dado que no se verán obligadas a aceptar más bienes primarios si no lo desean. Así, la premisa clave es que aun cuando las partes no posean suficiente información acerca de sus proyectos y fines particulares, tendrán suficiente conocimiento para jerarquizar las alternativas que tienen a su disposición. Guiadas las partes por esta teoría pragmática del bien, y por los hechos de la psicología moral, sus juicios no serán ya el mero producto de conjeturas superficiales. Pueden arribar a una decisión racional en el sentido ordinario de la palabra.

Una premisa especial de Rawls es que un individuo racional de la sociedad no es presa de la envidia; no le resulta a la persona insoportable el saber que otros tienen

una cantidad mayor de bienes primarios. Una razón para proceder de esta manera, un tanto idealista, es que la envidia tiende a hacer peores a todos los hombres. En este sentido la envidia es una desventaja para la totalidad de la sociedad. El autor supone que la ausencia de envidia equivale a la premisa de que, al elegir los principios, cada individuo pensará que tiene un plan de vida suficientemente valioso para sí.

Otro supuesto significativo es el de la racionalidad mutuamente desinteresada, que viene a decir que las personas en la posición original intentan reconocer principios de justicia que promuevan sus sistemas personales de fines. Lo hacen intentando ganar para sí la cantidad más grande posible de bienes sociales primarios, sin dañarse mutuamente, puesto que no están motivadas por la afección o por el rencor.

Existe una suposición adicional para asegurar la estricta observancia de las reglas reguladoras de la justicia. Se parte de que las partes son capaces de tener un sentido personal de la justicia, y que este hecho es de conocimiento público. Esta condición asegura la consistencia del cumplimiento del acuerdo hecho en la posición original.

Con estas observaciones acerca de la racionalidad y motivación de las partes representantes, queda suficientemente descrita la posición original, que puede ser resumida en los siguientes elementos que esquematizan su sustancia y sus límites.<sup>49</sup> Las alternativas con el subrayado constituyen la posición original de justicia como imparcialidad. Las otras opciones ilustran otras posibles posiciones iniciales.

1. La naturaleza de las partes:

- a. personas duraderas (jefes de familia o líneas genéticas);
- b. sujetos individuales;
- c. asociaciones (estados, iglesias, otras corporaciones).

2. Objeto de la justicia:

- a. estructura básica de la sociedad
- b. reglas para las corporaciones
- c. derecho internacional.

3. Presentación de alternativas

---

<sup>49</sup> *Ibidem.*, p.143

- a. lista más corta (o más larga);
  - b. caracterización general de las posibilidades.
- 4. Tiempo de ingreso:
  - a. cualquier tiempo (durante la mayoría de edad) para las personas vivas;
  - b. todas las personas reales (vivas en algún momento) simultáneamente;
  - c. todas las personas posibles simultáneamente.
- 5. Circunstancias de la justicia:
  - a. condiciones de escasez moderada;
  - b. las de *a* más otras extremas.
- 6. Condiciones formales para los principios:
  - a. generalidad, universalidad, carácter público, jerarquía y definitividad;
  - b. las de *a*, menos el carácter público.
- 7. Conocimiento y creencias:
  - a. velo de ignorancia;
  - b. información total;
  - c. conocimiento parcial.
- 8. Motivación de las partes:
  - a. desinterés mutuo (altruismo inmediato);
  - b. elementos de solidaridad social y de buena voluntad;
  - c. altruismo perfecto.
- 9. Racionalidad:
  - a. utilizar medios efectivos para alcanzar fines con expectativas unificadas y una interpretación objetiva de la probabilidad;
  - b. igual que *a*, pero sin expectativas unificadas y usando el principio de razón insuficiente.
- 10. La condición del acuerdo:
  - a. unanimidad perpetua;
  - b. aceptación mayoritaria o algo semejante por un período limitado.
- 11. La condición del cumplimiento:
  - a. cumplimiento estricto;
  - b. cumplimiento parcial en diversos grados.

12. El punto donde no hay acuerdo:

- a. egoísmo general;
- b. el estado de naturaleza.

Rawls profundiza en muchos de estos elementos de la posición original a lo largo de su obra. Por motivos de espacio, no me detendré a exponer las reflexiones que hace sobre cada uno de ellos. Baste de momento visualizar la metodología rawlsiana para la construcción de la situación hipotética inicial, momento definitorio y fundante para la elección de los principios de justicia.

## 17. LA PRESENTACIÓN DE OPCIONES

Rawls presenta varias alternativas de concepciones de justicia como accesibles a las partes representantes en la posición original. Pienso que este es un ejercicio muy amplio que da un gran calado a su propuesta pues plantea la elección de principios de justicia no sólo entre utilitarismo y justicia como equidad, sino que en este ejercicio abre las alternativas para poner a consideración de las partes una gran variedad de tradiciones.

El pensador admite que no puede afirmar que sus principios sean los mejores de todas las concepciones de justicia posibles, y postula que “incluso si es que existe una opción que sea la mejor, parece difícil describir las facultades intelectuales que deberían tener las partes para que conozcan este óptimo”.<sup>50</sup> Así, partiendo de que una optimización en el terreno de la justicia no es plausible, Rawls opta por un procedimiento de tipo práctico en el que genera una lista de las que él considera son las opciones más conocidas de justicia hasta ese momento, y asume que ésta les será presentada a las partes, quienes harán una comparación hasta que convengan de manera unánime qué concepción de justicia es la mejor. Podemos asumir que se llega a esta decisión a través de comparaciones en parejas.

Me parece relevante listar las opciones propuestas<sup>51</sup> en el texto:

---

<sup>50</sup> *Op. cit.*, p.123

<sup>51</sup> *Ibidem*, p.124

A. Los dos principios de la justicia (en orden serial):

1. El principio de mayor libertad equitativa;
2. a) El principio de justa igualdad de oportunidades;  
b) El principio de diferencia

B. Concepciones mixtas. Sustitúyase A(2) por uno de ellos:

1. El principio de utilidad media;
2. El principio de utilidad media, sujeto a restricciones, bien sea:
  - a) Que se mantenga un cierto mínimo social, o
  - b) Que la distribución no sea demasiado amplia, o

3. El principio de utilidad media sujeto a una de las restricciones en B(2) , así como a la  
de justa igualdad de oportunidades.

C. Concepciones teleológicas clásicas:

1. El principio clásico de utilidad;
2. El principio de utilidad media;
3. El principio de la perfección.

D. Concepciones intuicionistas

1. Equilibrar la utilidad total con el principio de igual distribución;
2. Equilibrar la utilidad media con el principio de la compensación;
3. Equilibrar una lista de principios *prima facie* (según requiera).

E. Concepciones egoístas<sup>52</sup>

1. Dictadura unipersonal: todos han de servir a mis intereses;
2. Privilegiado: todos han de actuar justamente, excepto yo, si así lo deseo;
3. General: todos pueden promover sus intereses en la medida en que les plazca.

El autor especifica una mecánica concreta para argumentar que sus dos principios de justicia son superiores a todas las alternativas presentadas: primero se concentra en avalar la superioridad frente al principio de utilidad, y más adelante refuta de manera

---

<sup>52</sup> Ver sección 23 de *Teoría de la Justicia* para la explicación de Rawls de por qué esas concepciones egoístas no son opciones.

general las otras variantes, a partir de la argumentación contra el utilitarismo. Expondré en el apartado 18, de manera sucinta, la argumentación relativa al utilitarismo medio, que ilustra con claridad el razonamiento rawlsiano, a fin de permanecer dentro del alcance establecido para esta investigación.

En la siguiente sección presentaré las restricciones propuestas por Rawls, y sus respectivas argumentaciones, para servir como acotamientos a las partes de manera que puedan definir criterios de selección para elegir los mejores principios.

## **18. LAS RESTRICCIONES FORMALES DEL CONCEPTO DE LO JUSTO**

Los principios de justicia a ser elegidos deben cumplir primordialmente la misión de resolver las demandas que las personas se hacen unas a otras, y a sus instituciones. Para reflejar este atributo cualitativo, Rawls establece cinco condiciones que deberán cumplir los principios:

En primer lugar, deberán ser generales. Esto es, los predicados utilizados en su enunciación deberán expresar relaciones generales, y no usar palabras que puedan ser reconocidas como nombres propios o definiciones concretas. La premisa es que, dado que las partes no disponen de información particular de su situación, no serán capaces de elaborar principios que les sean ventajosos. Así, las partes se ven forzosamente obligadas a acordar principios generales. La naturalidad de esta condición se funda en el hecho de que los primeros principios deben poder servir como base pública permanente de una sociedad bien ordenada.

Una segunda restricción es que tendrán que ser universales en su aplicación. En otras palabras, los principios deben ser válidos para todos los miembros de la sociedad, por ser personas morales. La idea es que cualquier individuo de capacidad intelectual elemental pueda entenderlos y hacer uso de ellos en sus reflexiones personales o públicas. Esto implica que la complejidad no será muy elevada, y que la cantidad de distinciones que hacen no sea tampoco excesiva.

En tercer lugar, los principios serán de carácter público, puesto que las partes asumen que están eligiendo enunciados para una concepción pública de justicia, congruente con la teoría del contrato social. Se presume que todos los miembros de la comunidad sabrán todo lo concerniente a estos principios. Debemos tener cuidado en

distinguir el carácter público con la universalidad de su aplicación, pues esta última refiere a que los principios serán observados por todos de manera consciente y consistente, mientras que la condición de publicidad busca que las partes sopesen las concepciones de la justicia como parte integral de la vida social cotidiana.

Una restricción adicional es que una concepción de lo justo debe establecer una ordenación de los requerimientos ciudadanos que puedan ser conflictivos. Esta exigencia surge de manera natural del papel que tendrán sus principios para conciliar las peticiones encontradas. El desafío es decidir lo que ha de considerarse como una ordenación, que idealmente sería completa, de manera que fuera capaz de ordenar la totalidad de las demandas que puedan surgir. Idealmente, esta ordenación debería también ser transitiva.

La quinta y última condición de los principios es su carácter definitivo. Las partes deberán valorar al sistema de principios como un tribunal de máxima autoridad en apelación a temas de razonamiento práctico sobre la justicia. No hay recursos más elevados a los que pueda recurrirse para dirimir diferencias y, por tanto, los razonamientos derivados de los principios deben ser concluyentes.

Pienso que Rawls apela a estas cinco condiciones con dos objetivos. Primeramente, para establecer unos niveles mínimos de calidad y de funcionalidad social deseable en los principios, que refuerzan la premisa idealista de una sociedad bien ordenada. Pero también argumento que resulta sospechoso que Rawls acuda a este artificio de los condicionantes, pues claramente se podría interpretar como la integración al proceso de elección de principios, de unos criterios que favorecerán a los dos principios de Rawls. En este sentido, me parece que las 'restricciones formales del concepto de lo justo' resultan cuestionables.

## **19. RAZONAMIENTO QUE CONDUCE A LOS DOS PRINCIPIOS DE JUSTICIA**

El reto más importante del sistema de Rawls es determinar la superioridad de sus dos principios sobre el de la utilidad, en base a un razonamiento estrictamente racional.

Primero revisaré las observaciones intuitivas que hace el autor a favor de las dos alternativas, a partir del punto de vista de una persona que se encuentra situada en la posición original. Debido al velo de ignorancia, no hay forma de que esta persona

obtenga alguna ventaja especial, pero tampoco tiene motivos para aceptar desventajas. Así, su postura inicial más racional sería requerir una división perfectamente igualitaria de los bienes disponibles: libertades, oportunidades e ingresos y riqueza.

Sin embargo, aun cuando la persona en la situación original se mantenga firme respecto a la importancia de la igualdad de libertades y de oportunidades, Rawls postula, me parece con razón, que las fuerzas intrínsecas de la economía y las necesidades de una eficiente organización provocarán desigualdades en el ingreso y en el grado de responsabilidad, que harán que todos estén mejor que en una situación hipotética de igualdad total. Partiendo de que las partes no están nunca presas de la envidia, se permitirán estas desigualdades con la condición de que todos mejoren su nivel de ingresos y riqueza, siempre y cuando las diferencias vayan acompañadas de una distribución equitativa de las oportunidades y una libertad igual. Aquí Rawls otorga a los menos favorecidos un derecho de veto, que se expresa en el principio de diferencia: aquellos que se han beneficiado más, deberán haberlo hecho en condiciones que sean justificables para aquellos que han ganado menos. Así, por medio de este razonamiento las partes representantes pueden llegar a los dos principios de justicia rawlsianos serialmente ordenados.

La defensa de los dos principios se solidifica a través de la introducción de la noción de persona libre, la cual tiene varios rasgos: en primer lugar, dan una elevada prioridad a la manera en que sus intereses, incluidos los fundamentales, serán regulados por las instituciones a establecer. Otra característica es que no se consideran como forzosamente obligadas a perseguir un conjunto específico de intereses fundamentales, aunque desean tener la facultad de promover dichos intereses. Especulo que aquí hay una alusión al utilitarismo y su valoración exclusiva a la maximización de la utilidad; la persona libre no se percibe comprometida con esa causa única. En contraste, la persona libre se considera con la facultad de revisar y alterar sus objetivos últimos, y da una prioridad a la vigencia de sus libertades fundamentales. Dado que los dos principios aseguran una estructura social consistente con esta condición, argumenta Rawls, éstos serán preferidos por los hombres, por encima del principio de la utilidad. Nuevamente encuentro evidencias de la introducción de nociones, como esta de persona libre, que hacen las veces de tamiz para que los principios del autor sean

elegidos. Claramente, si la 'persona libre' no puede centrarse en un solo objetivo, como es el caso de la utilidad, los principios del utilitarismo quedarán automáticamente descartados.

El autor enfatiza que la prioridad de la libertad significa que no será posible intercambiar una libertad menor por una mejora en el nivel económico. De hecho, sólo en la condición extrema de una necesidad imperiosa de cambiar las condiciones de la civilización será factible la negación de las libertades iguales, de forma que en un tiempo establecido se puedan reinstaurar libertades equitativas. Hasta aquí, los dos principios constituyen, al menos, una propuesta plausible de justicia; sin embargo es necesario postular una argumentación más sistemática a su favor. Rawls propone estudiar las consecuencias de estos dos principios en las instituciones y compararlas con nuestros juicios intuitivos acerca de lo que debiera ser la justicia, tema que se cubre en la segunda parte del libro, que no revisaré en detalle en este trabajo. El otro argumento central a favor de los principios, es que son "la solución maximin al problema de la justicia social".<sup>53</sup> En este sentido, "los dos principios son aquellos que escogería una persona al proyectar una sociedad en la cual su enemigo hubiera de asignarle su lugar."<sup>54</sup> Como expondré más adelante, este argumento maximin para la defensa de los dos principios es el foco del ataque de Harsanyi en su crítica.

Rawls aclara que las personas en la posición original no asumen que su lugar en la sociedad será determinado por un enemigo con malas intenciones, pero no modifica la condición maximin, que parte precisamente de esta suposición. Un argumento adicional a favor de maximin es que en la posición original será racional para las partes adoptar una posición conservadora para protegerse contra la contingencia del caso del peor enemigo.

Aunque Rawls indica que "es claro que la regla maximin no es, en general, una guía apropiada para elegir bajo circunstancias de incertidumbre"<sup>55</sup>, el pensador provee otro argumento a favor de este criterio al proponer que la posición original exhibe en alto

---

<sup>53</sup> *Op. cit.*, p.150.

<sup>54</sup> *Idem*

<sup>55</sup> *Ibidem*, p.151

grado tres rasgos que hacen atractiva esta regla. Primero, dado que maximin, no toma en cuenta las probabilidades, la posición original deberá ser tal, que no sea posible estimar ninguna probabilidad. Aquí entra en acción el velo de ignorancia, haciendo que las partes no tengan la facultad de asignar ningún tipo de probabilidad a los posibles arreglos sociales ni al lugar que ellas ocuparán en la escalera social a establecer. El segundo rasgo es que “la persona que escoge tiene una concepción del bien tal que le importa muy poco o nada lo que pueda ganar todavía por encima del mínimo estipendo que seguramente obtiene al seguir la regla maximin.”<sup>56</sup> En otras palabras, no es prudente para la persona correr el riesgo de quedar por abajo del nivel asegurado por la regla, en la expectativa de terminar por encima de ese valor. El tercer rasgo de la posición original es que “las alternativas rechazadas tienen resultados que difícilmente pueden aceptarse”<sup>57</sup>, puesto que la situación involucra riesgos elevados en el sentido de que los principios a establecerse serán permanentes y no será posible modificarlos posteriormente para buscar terminar en una posición mejor a la que nos fue asignada. Las alternativas rechazadas incluyen obviamente al utilitarismo, el cual no sería aceptable por su frecuente disminución de libertades individuales a cambio de mayores beneficios sociales.

También señala el autor que el utilitarismo depende en mayor grado de argumentos basados en hechos generales, en lugar de modelar condiciones morales concretas, pues tiende a refutar las objeciones argumentando que las leyes sociales y de la naturaleza humana eliminan de una manera ‘natural’ aquellos casos de injusticia que se considerarían intolerables de acuerdo con nuestros juicios meditados sobre la justicia. Por el contrario, justicia como imparcialidad incorpora de una manera directa los ideales de la justicia como se les concibe comúnmente. Hay dos razones que a los ojos de Rawls justifican la incorporación de estos ideales en su sistema. En primer lugar, los argumentos del utilitario para suponer que las limitaciones a la libertad serán rara vez justificables son “sólo probablemente verdaderos, e incluso dudosos.”<sup>58</sup> Por

---

<sup>56</sup> *Idem*

<sup>57</sup> *Idem*

<sup>58</sup> *Ibidem*, p.156

ello, las partes optarán por asegurar sus libertades en vez de hacerlas depender de cálculos numéricos especulativos. En segundo lugar, es altamente deseable evitar los complicados argumentos teóricos del utilitarismo para llegar a una concepción pública de justicia. Así, puesto que en justicia como imparcialidad las nociones morales son difundidas públicamente, los beneficios de esta declaración pública mueven la balanza a favor de los dos principios, aun cuando las premisas del utilitarismo fuesen realmente verdaderas.

Rawls complementa su defensa de los dos principios con una crítica al razonamiento que lleva al principio de la utilidad. El argumento central del autor es que en la posición original las personas tendrán una muy elevada aversión al riesgo, debido a las enormes implicaciones de la elección a llevarse a cabo, y a la incertidumbre total generada por el velo de ignorancia. De hecho, a los ojos de Rawls, la aversión al riesgo en la situación inicial puede ser tan alta que la ponderación del método utilitario lleve a un resultado práctico tan cercano al principio de diferencia que la simplicidad de este último lo haga más favorable como criterio de justicia. Este argumento está basado en un criterio técnico que está fuera del alcance de esta investigación.

En sinopsis, Rawls razona a favor de su sistema, y en contra del utilitarismo, a partir de varios argumentos y premisas fundamentales. Los menos aventajados tendrán derecho de veto para validar que aquellos que han ganado más estén justificados; la condición será la que establece el principio de diferencia: las desigualdades serán permitidas solamente si todos, incluyendo los menos favorecidos, mejoran su ingreso. Las libertades tendrán prioridad sobre la distribución de ingreso y riqueza; el utilitarismo no puede garantizar este orden serial, y justicia como imparcialidad sí lo logra. Los dos principios son la solución maximin que permite una protección efectiva contra la contingencia de que el nuestro lugar sea asignado por el enemigo. Por otro lado, la 'persona libre', no se verá a sí misma obligada a perseguir un solo interés por encima de sus libertades y de sus otros intereses fundamentales, como es el caso del utilitarismo. Finalmente, la posición original exhibe en alto grado tres rasgos que hacen conveniente la aplicación de maximin: no es posible el uso de probabilidades, a las personas les importa muy poco el bienestar que pudieran tener por encima del

estipendio garantizado por maximin y, por último, las alternativas rechazadas, incluido el utilitarismo, serían difícilmente aceptables.

Queda así descrito el método general de razonamiento de Rawls para demostrar la superioridad de sus principios sobre el de la utilidad media. Es sabido que este ejercicio ha sido sujeto de muchas críticas, entre las cuales está la que me ocupa en este trabajo de investigación.

## **20. EXPRESIÓN DEFINITIVA DE LOS DOS PRINCIPIOS DE JUSTICIA**

En la sección 13 presenté la formulación preliminar de los principios de justicia como imparcialidad. Una vez que Rawls ha elaborado sus argumentos sobre la preferencia por sus principios en el contexto de las instituciones y de la política social, el autor expone su enunciación definitiva<sup>59</sup> para la sociedad:

### *Primer Principio*

Cada persona ha de tener un derecho igual al más extenso sistema total de libertades básicas compatible con un sistema similar de libertad para todos.

### *Segundo Principio*

Las desigualdades económicas y sociales han de ser estructuradas de manera que sean para:

- a) mayor beneficio de los menos aventajados, de acuerdo con un principio de ahorro justo, y
- b) unidos a los cargos y las funciones asequibles a todos, en condiciones de justa igualdad de oportunidades.

Comparando con la versión preliminar, podemos ver claramente que, mientras el primer principio es prácticamente similar a su versión anterior, el segundo principio ha sido refinado, al establecer que las desigualdades sociales y económicas -que anteriormente debían conformarse de manera que fuesen “ventajosas para todos”-

---

<sup>59</sup> *Op. cit.*, p.280

ahora lo serán de forma que sean para el “mayor beneficio de los menos aventajados”. En este punto Rawls formaliza el principio de diferencia tal como se le conoce y como se le ha criticado.

## **21. LA MORALIDAD RAWLSIANA**

El desarrollo de la teoría de justicia sólo es posible a través de un procedimiento que permita la elección racional de unos principios que puedan ser moralmente aceptables. Ya expuse el razonamiento, descrito como un proceso ‘racional’, para la elección de los principios. En este apartado expondré la forma en que Rawls desarrolla la motivación moral para la validación de los principios de justicia como moralmente correctos.

Uno de los principales rasgos de la explicación del desarrollo moral rawlsiano es la idea de reciprocidad, entendida como “un profundo hecho psicológico”<sup>60</sup>, por medio del cual formamos lazos afectivos a personas e instituciones según la forma en que percibamos que nuestro bienestar es afectado por ellos. En base a esta tendencia a la reciprocidad Rawls formula tres leyes psicológicas<sup>61</sup> de desarrollo moral, y las aplica para argumentar que normalmente las personas buscarán hacer justicia en respuesta a haber sido tratados con justicia. Además, bajo las circunstancias especiales de una ‘sociedad bien ordenada’, las personas eventualmente desarrollan un deseo de apoyar a las instituciones justas dado que éstas los benefician a ellos y a aquellos que les importan.

Los principios de reciprocidad de Rawls cubren tres etapas de desarrollo moral de la persona. La primera ley dice en la primera etapa de progreso moral, cuando la institución familiar es ella misma justa, el niño gradualmente desarrolla lazos amorosos por sus padres cuando ellos lo han cuidado amorosamente. Con el tiempo, el amor y la confianza de los padres llevan al hijo a adherirse a las órdenes de los padres, mas no tanto por temor a una sanción, sino como producto del amor y el respeto por los padres y por el deseo de emularlos y no decepcionarlos. Rawls denomina a esta etapa la ‘moralidad de la autoridad’.

---

<sup>60</sup> *Ibidem*, p.447

<sup>61</sup> *Ibidem*, p.443

La segunda etapa de crecimiento moral, o 'moralidad de la asociación' es caracterizada por el aprendizaje y los lazos emocionales dirigidos a estándares morales que son apropiados para nuestro rol social, como puede ser compañero de clase, o amigo, en las asociaciones varias a las que pertenecemos. A partir de estas asociaciones el muchacho aprende las virtudes de carácter propias del rol que se tiene en cada asociación. Uno de los rasgos más relevantes de esta etapa es que, a medida que el muchacho alcanza a comprender los diferentes roles y posiciones que constituyen los esquemas de asociación, él desarrolla una concepción del sistema de cooperación que define a la asociación y los propósitos que ésta promueve.

Por último, la tercera etapa del crecimiento moral es la 'moralidad de principios'. Asumamos que una persona ha experimentado actitudes de amor y confianza dentro de la familia, y de confianza mutua dentro de otras asociaciones, en concordancia con las dos primeras leyes psicológicas. Rawls argumenta que desarrollamos un sentido de justicia, que implica el deseo de apoyar a las instituciones justas, una vez que reconocemos que nosotros y aquellos que nos son importantes nos beneficiamos de las instituciones justas, una vez que estas instituciones son públicamente reconocidas como justas. Esta motivación moral se muestra en nuestra disposición a cumplir con nuestro deberes y obligaciones, y a hacer la parte que nos corresponde en instituciones justas, y en nuestro deseo de reformarlas cuando llegan a ser injustas.

Más adelante<sup>62</sup>, Rawls argumenta que los principios de justicia de su teoría son compatibles con los tres principios de reciprocidad y que, por tanto, los miembros de una sociedad bien ordenada generalmente desarrollarán un sentido de la justicia adecuado para actuar según estos principios. El sentido de justicia es un elemento clave para lograr lo que Rawls denomina 'estabilidad relativa'; su objetivo es mostrar que su teoría es más estable que las otras concepciones de justicia consideradas, con claro énfasis en la superioridad al ser comparada con el utilitarismo.

Así, Rawls construye su teoría de moralidad, y argumenta que las partes poseerán un sentido de justicia tal, que les permitirá ejercer un juicio moral que reconozca que

---

<sup>62</sup> Ver capítulo 8 de *Teoría de la Justicia*.

los dos principios de justicia racionalmente aceptados son también los mejores desde el punto de vista moral.

## 22. INFLUENCIA DEL CONTRACTUALISMO DE LOCKE Y ROUSSEAU

Podemos decir que Rawls es el proponente por excelencia de la teoría del contrato social del siglo XX. Sin embargo, el autor reconoce a sus antecesores cuando dice “lo que he tratado de hacer es generalizar y llevar la teoría tradicional del contrato social representada por Locke, Rousseau...a un nivel más elevado de abstracción.”<sup>63</sup> A continuación expongo brevemente las visiones de Locke y Rousseau sobre este tema.

Mientras que para de Hobbes, como ya comenté en el primer capítulo, la necesidad de una autoridad en la forma del soberano surge de la brutalidad del estado de la naturaleza, para John Locke esta situación natural inicial es un lugar muy distinto, y por tanto sus argumentos acerca del contrato social y de la naturaleza de la relación del hombre con la autoridad son muy diferentes. Los textos más importantes e influyentes de Locke están contenidos en sus *Dos Tratados sobre el Gobierno*<sup>64</sup>. El primer tratado se ocupa básicamente de la refutación del argumento de *El Patriarca*, de Robert Filmer, cuya tesis era que la autoridad política se deriva de la autoridad religiosa, una especie de derecho divino de los reyes. El segundo tratado incluye la visión de Locke acerca del objetivo y la justificación de un gobierno civil, bajo el nombre *Un Ensayo sobre el Verdadero Alcance y Fin de un Gobierno Civil*.

Para Locke, el Estado de la Naturaleza es un estado perfecto y de libertad completa para llevar a cabo nuestros fines según lo desee cada quien, libre de la interferencia de los otros. Pero esto no significa que cada quien pueda hacer lo que le plazca, pues existe un sentido de moralidad. El Estado de la Naturaleza es pre-político, mas no es pre-moral. Así, se asume que las personas son iguales y por tanto igualmente capaces de estar ligados por la Ley de la Naturaleza. Esta ley mayor, que para Locke soporta toda la moralidad, fue dada al hombre por Dios, y exige que no hagamos daño a los otros en relación a la vida, la salud, la libertad, y las posesiones. Dado que todos

---

<sup>63</sup> *Op. cit.*, p.10.

<sup>64</sup> John Locke, *Dos Tractos Sobre el Gobierno*, Tr. Daniel Mielgo. Madrid: Biblioteca Nueva, 2015.

pertenece a Dios por igual, y que no podemos tomar lo que es de Él, tenemos prohibido hacernos daño entre nosotros humanos.

El Estado de la Naturaleza no es un estado de guerra como lo concibe Hobbes, pero sí puede convertirse en un estado de guerra, particularmente por las disputas sobre la propiedad. Puesto que el Estado de la Naturaleza no tiene autoridad civil, una vez que comienza la guerra, es muy probable que continúe; y es por ello que el hombre modifica el Estado de la Naturaleza al hacer un contrato colectivo para formar el gobierno civil.

Para Locke, el estado inicial no es una condición de los individuos. Más bien, es una sociedad poblada por familias -por madres y padres con sus hijos. Estas sociedades están basadas en acuerdos voluntarios para cuidar de los hijos, y son morales, mas no políticos. La sociedad política se constituye cuando de manera individual los hombres, representantes de sus familias, se reúnen en el Estado de la Naturaleza y acuerdan ceder el poder ejecutivo para castigar a aquellos que transgredan la Ley Natural, y entregan ese poder al poder público de un gobierno. Al crear la sociedad política y el gobierno, el hombre gana tres cosas que no tenía en el Estado de la Naturaleza: leyes, jueces para aplicar las leyes, y el poder ejecutivo para imponer esas leyes.

Cuando el poder político del gobierno deviene en tiranía, y niega a la población la capacidad de hacer leyes para su propia preservación, el tirano se pone en un estado de guerra con los gobernados y ellos tienen el derecho de autodefensa y de resistir la autoridad del gobernante. Así, el contrato político se disuelve y empieza nuevamente un proceso para crear la siguiente sociedad política.

Voy ahora con Jean Jacques Rousseau, quien nace ochenta años después de Locke, en el siglo XVIII de la ilustración, en una Francia de gran actividad intelectual. Rousseau tenía dos teorías del contrato social. La primera se encuentra en su *Discurso sobre el Origen de la Desigualdad entre los Hombres*<sup>65</sup>, y da cuenta de la evolución moral y política del ser humano en el tiempo, desde un Estado de la Naturaleza hasta la sociedad moderna. Como tal, la historia incluye el desarrollo del contrato social, que

---

<sup>65</sup> Jean-Jacques Rousseau, *Discurso Sobre el Origen de la Desigualdad entre los Hombres*, Tr. Fito Rodríguez. Madrid: Delta, 2013.

para él es muy problemático. La segunda obra, *Del Contrato Social*<sup>66</sup>, es su teoría normativa idealista del contrato social, y tiene como objetivo brindar los medios para aliviar los problemas que la sociedad moderna ha creado para los hombres.

Al describir el proceso histórico por el que el hombre 'progresa' al pasar de un Estado de la Naturaleza a una sociedad civil, Rousseau tiene la visión de un estado inicial pacífico y agradable en el que las personas vivían de manera solitaria y poco complicada. Las pocas necesidades eran fácilmente satisfechas por la naturaleza. En medio de la abundancia y debido a la poca población, había poco contacto entre la gente y no había razón para tener conflictos ni temores. Estas personas moralmente puras tenían la capacidad natural de la misericordia y por tanto no tenían inclinación de hacer daño a los demás.

Con el paso del tiempo, la humanidad enfrentó ciertos retos. Con el aumento de la población, los medios que usaba la gente para satisfacer sus necesidades tuvieron que cambiar. Las personas poco a poco empezaron a vivir juntas en pequeñas familias, y después en pequeñas comunidades. La división de labores fue introducida y los descubrimientos fueron haciendo la vida más fácil, surgiendo así el tiempo libre. La disponibilidad de tiempo libre fue inevitablemente llevando a las comparaciones, generando así los valores públicamente reconocidos, la vergüenza y la envidia, y el orgullo y el desprecio. Pero lo más importante, para Rousseau, fue la invención de la propiedad privada, que generó la avaricia, la competencia, la vanidad, la desigualdad y el vicio. Para Rousseau la invención de la propiedad constituye la caída de la humanidad de la gracia de la Naturaleza.

Eventualmente, aquellos con propiedades caen en la cuenta de que les conviene crear un gobierno que proteja la propiedad privada de aquellos que no la tienen, pero que piensan que podrían tenerla por el uso de la fuerza. Así, el gobierno se establece, a través de un contrato, con el propósito de garantizar la igualdad y la protección de todos, aunque su verdadero propósito es perpetuar las desigualdades que la propiedad privada ha producido. En otras palabras, el propio contrato que afirma existir para igual beneficio de todos, es en realidad un vehículo para favorecer a aquellos que se han

---

<sup>66</sup> Jean-Jacques Rousseau, *Del Contrato Social*, Tr. Mauro Armíño. Madrid: Alianza, 2012.

vuelto más fuertes y más ricos como resultado del desarrollo de la propiedad privada. Este es el contrato social naturalizado, el cual Rousseau ve como el responsable del conflicto y la competencia que se sufre en las sociedades modernas.

El contrato social normativo propuesto por Rousseau en *Del Contrato Social*, tiene la misión de responder a este desafortunado estado de cosas, y remediar los males sociales y morales que se han producido en el desarrollo de la sociedad. La distinción entre la actual situación fáctica del hombre, y el cómo debería de vivir es muy importante para Rousseau. Aunque no debemos ignorar la historia, debemos resolver los problemas sociales a través de nuestra capacidad para elegir cómo debemos vivir.

Rousseau escribe: “El hombre nace libre, pero en todos lados está encadenado.”<sup>67</sup> Los seres humanos son esencialmente libres, y eran libres en el Estado de la Naturaleza, pero el ‘progreso’ de la civilización ha sustituido la libertad por la servidumbre hacia otros a través de la dependencia, las desigualdades económicas y sociales, y el grado en que nos juzgamos a nosotros mismos a partir de la comparación con los otros. Puesto que el regreso al Estado de la Naturaleza no es posible ni deseable, el propósito de la política es restablecer la libertad para el hombre, reconciliando así lo que realmente somos con la manera en que vivimos juntos. Así, la pregunta fundamental que la obra busca responder es: ¿cómo podemos ser libres y vivir juntos? En otras palabras, ¿cómo podemos vivir juntos sin sucumbir a la fuerza y a la coerción del otro? Podemos hacerlo, sostiene Rousseau, subordinando nuestra voluntad individual particular a la voluntad colectiva, creada a través del contrato con otras personas libres e iguales. El soberano es entonces formado cuando personas iguales y libres se reúnen y acuerdan crearse de nuevo como una sola entidad, a ser dirigida para el bien de todos.

Para Rousseau, esto implica una forma de democracia muy fuerte y directa. El individuo no puede transferir a otro su voluntad, como en las democracias representativas. Más bien, la voluntad general se activa en la reunión colectiva periódica del cuerpo democrático completo formado por todos los ciudadanos, para decidir en conjunto, con condición cercana a la unanimidad, cómo vivir juntos, y qué

---

<sup>67</sup> *Ibidem*, p. 49.

leyes aceptar. Pienso que esto implicaría que esta forma fuerte de democracia sólo sería posible en estados relativamente pequeños en los que todos se conozcan entre sí. Aunque estas condiciones democráticas son muy estrictas, para el autor ésta es la única manera en la que la humanidad puede salvarse a sí misma y recuperar la libertad a la que cada individuo tiene derecho de manera natural.

Me parece que las teorías del contrato social de Rousseau en su conjunto forman una visión consistente de nuestra situación moral y política. El ser humano es dotado con libertad e igualdad por la naturaleza, pero su naturaleza ha sido corrompida por su historia social contingente. Sin embargo, podemos superar esta corrupción si invocamos a nuestra voluntad libre para reconstituircos políticamente bajo principios democráticos fuertes que sean buenos para el ser humano, tanto en lo individual como en lo colectivo.

Así, Rawls retoma la tradición contractual de estos dos importantes pensadores, y agrega nuevos elementos, algunos propios y otros de doctrinas diferentes, para construir su edificio conceptual.

### **23. TRES CRÍTICAS A LA TEORÍA DE JUSTICIA DE RAWLS**

Como es sabido, Rawls es considerado por algunos pensadores académicos como el filósofo más importante del siglo XX, dado el alcance y la originalidad de su Teoría de Justicia, publicada en 1971, y los subsecuentes esfuerzos de refinación a la misma, dados a conocer en publicaciones posteriores.

Sin embargo, la propuesta rawlsiana ha sido sujeto de numerosas críticas por parte de académicos del más alto calibre. En este apartado revisaré las objeciones postuladas por Michael Sandel, filósofo conservador de la escuela comunitarista; las discrepancias presentadas por el filósofo y experto en teoría de decisión racional Brian Barry; y los inconvenientes expuestos por el reconocido pensador de la tradición liberal –en cierto sentido “compañero” de nuestro autor- Jürgen Habermas.

Crítica de Sandel

En su libro *El Liberalismo y los Límites de la Justicia*<sup>68</sup>, Michael Sandel hace una fuerte reprobación a la Teoría de Justicia de Rawls. La crítica de Sandel se basa en su argumento de que “lo que emana en un extremo en una teoría de justicia, debe emanar en el otro en una teoría de la persona, o más precisamente, en una teoría de la persona moral”<sup>69</sup>. Sandel argumenta que la teoría de justicia de Rawls requiere que el sujeto moral sea un agente abstracto de elección, completamente separado de sus fines, atributos personales, comunidad o historia. Sólo adoptando esta noción de persona de extrema individualidad hace sentido la teoría de justicia de Rawls. Esta individualidad radical no permite que la comunidad sea quien constituya a la persona, y tampoco permite la posibilidad de que la identidad significativa de la persona sea más un tema cognitivo que de elección.

La crítica de Sandel consta de dos partes. En la primera Sandel argumenta que la teoría de la persona autónoma postulada por Rawls es inconsistente con su principio de diferencia. En efecto, el principio de diferencia requiere que las instituciones básicas de la sociedad maximicen la posición de los menos favorecidos. Sandel afirma que, si el sujeto moral es un individuo, entonces el principio de diferencia implica el retiro de los talentos de unas personas para beneficiar a los menos favorecidos; por tanto, el principio de diferencia trata a esos sujetos como medios y no como fines. Sólo un grupo, o sujeto comunidad, podría simultáneamente elegir el principio de diferencia y, dado que los talentos de las personas pertenecerían al grupo, evitar al mismo tiempo tratar al sujeto moral como medio. Por tanto, en vista de que la teoría rawlsiana del sujeto moral es incongruente con el principio de diferencia, su noción de persona individualizada no es congruente con su teoría de justicia.

La segunda crítica de Sandel refiere a que Rawls está comprometido con una noción débil o estrecha de la persona –una persona separada de sus fines, compromisos y capacidades. Esta persona es tan escasa que no puede constituirse en un objeto de auto-reflexión. Sólo puede ser un sujeto, cuando más, capaz de elecciones arbitrarias y básicamente sin sentido. La arbitrariedad y el sin sentido de

---

<sup>68</sup> Michael Sandel, *El Liberalismo y los Límites de la Justicia*, Tr. María Luz Melón. Barcelona: Gidesa, 2009.

<sup>69</sup> *Ibidem*, nota 1, p.47

estas elecciones producen otra falla: una teoría inadecuada de lo bueno. Estas dos críticas en conjunto argumentan que la noción de persona rawlsiana no es atractiva, ni consistente con nuestra comprensión y experiencia de nosotros mismos, ni adecuada para respaldar la teoría de justicia de Rawls. En efecto, escribe Sandel:

La concepción de Rawls de la persona no puede respaldar su teoría de justicia, y dar cuenta de una manera creíble sobre nuestras capacidades de agenciamiento y auto reflexión; la justicia no puede ser primaria en la manera que la requiere la deontología, puesto que no podemos vernos a nosotros mismos coherentemente como el tipo de seres que la ética deontológica requiere que seamos.<sup>70</sup>

Más específicamente, la teoría de Rawls es inconsistente con seres que están constituidos por sus valores, carácter, compromisos y prácticas, que son parcialmente constituidos por ser miembros participantes de una comunidad, y dedicados a la auto-reflexión. En la introducción a este trabajo ya presenté mi conclusión personal sobre la antropología de Rawls, y coincido con Sandel en afirmar que su noción de hombre es vacía.

### Crítica de Barry

El filósofo inglés Brian Barry, reconocido experto en teoría de elección social, publicó una severa crítica<sup>71</sup> a la teoría rawlsiana en 1973, tan sólo dos años después de que *Teoría de la Justicia* se diera a conocer.

La crítica es muy extensa y profunda, pero pienso que puede ser esquematizada en tres elementos. Un primer tema de la crítica de Barry refiere a los esfuerzos de Rawls, que el inglés considera intentos fallidos, de mantener unas premisas ‘débiles’ –al buscar que sean independientes de alguna concepción particular del bien o del ideal de persona- para alcanzar conclusiones ‘fuertes’; en opinión de Barry esta deficiencia con frecuencia lleva al pensador americano a una situación sin salida. En efecto, el resultado, de acuerdo con Barry, es uno de estos tres: o i) las conclusiones de Rawls

---

<sup>70</sup> *Idem.*, p. 65

<sup>71</sup> Brian Barry, *The Liberal Theory of Justice, A Critical Examination of the Principal Doctrines in A Theory of Justice by John Rawls*. Londres: Oxford University Press, 1973.

no son lo que él ha planteado; o ii) sus premisas no son lo que él pretende; o iii) los argumentos esgrimidos no son válidos. Barry hincó el diente agudamente en toda la argumentación relevante de Rawls, particularmente en la derivación de los bienes primarios, la derivación de la prioridad de la libertad, la derivación de la igualdad de oportunidad, la derivación del criterio maximin, y el uso de maximin en teoría social. En pocas palabras, Barry cuestiona la solidez de todos los elementos de la estructura rawlsiana.

Barry concluye que la única salida a este “trilema” es que los liberales retiren la premisa rawlsiana fundamental de igualdad y acepten que hay tipos de caracteres que son más valiosos y admirables que otros, y que por tanto la sociedad debiera estar organizada para incentivar a los individuos más importantes a trabajar y producir para el beneficio de todos, incluyendo -más no de manera prioritaria- el de las personas menos favorecidas.

Expongo brevemente una de las críticas para ilustrar el razonamiento de Barry. En el caso de la posición original, el momento fundante de la teoría rawlsiana, Barry reprueba la ‘deducción lógica’, pretendida por Rawls, de los principios de justicia a partir de las condiciones de la condición original. El inglés argumenta que, en realidad, la labor de Rawls es ajustar las circunstancias de la situación original de manera tal, que solamente los principios rawlsianos puedan ser seleccionados. Dice Barry:

La afirmación (de Rawls) de haber establecido una relación deductiva...sería de esta forma: si los principios de justicia son aquellos principios que serían elegidos en una posición original caracterizada como la caracteriza Rawls, entonces los ‘dos principios’ son los principios de justicia.<sup>72</sup>

Otro de las críticas de Barry es que el tipo de liberalismo de Rawls es, en algunos aspectos, arcaico. Un primer ejemplo de ello es el que Rawls considere que el auto-respeto se asegura con la garantía de igualdad de derechos y oportunidades, y el que la cuestión de si este auto-respeto es afectado por la porción de riqueza y poder del individuo ni siquiera es abordada. Otro ejemplo refiere a la definición de Rawls de los menos favorecidos como aquellos que tienen menos ingresos y riqueza; esta

---

<sup>72</sup> *Ibidem*, p.23

esquematación ignora lo que para Barry es una tesis “actualmente aceptada” de que la elevación del ingreso promedio de los trabajadores no calificados no ataca a las verdaderas causas de la pobreza. Es decir que la definición de Rawls a través de la sola dimensión del ingreso no puede ser suficiente para capturar los rasgos de condición de los menos favorecidos.

Un tercer tema de la reprobación es que la crítica de Rawls contra el utilitarismo es injusta en varios aspectos. Por ejemplo, Barry considera inválidos muchos de los argumentos rawlsianos, relativos a que los principios del utilitarismo no serían aceptados en la posición original; y a sus ojos deberían descartarse o ser reformulados significativamente. Por otra parte, dice Barry, Rawls critica al utilitarismo por requerir premisas empíricas especiales, cuando su propia teoría requiere de una serie de premisas desconectadas de la realidad, como, por ejemplo, el principio Aristotélico de que por racionalidad las personas preferirán dedicarse, entre dos actividades, a aquella que les exija usar un mayor repertorio de habilidades adquiridas. En la visión de Barry, este es un intento por parte de Rawls de introducir una idea substancial de la excelencia humana bajo la sombrilla –aparentemente neutral- de la racionalidad.

Me parece que la mayoría de los argumentos de Barry son dirigidos a la falta de consistencia interna de la propuesta de Rawls; es decir, busca mostrar cómo la propia argumentación de Rawls no se sostiene desde el punto de vista estrictamente lógico racional, al no haber una conexión sólida entre premisas y conclusiones. Por lo tanto, concluye Barry, una teoría con esas fallas no puede funcionar. Esto es evidente cuando el autor inglés escribe al inicio del texto: “El rumbo general de este libro, como el lector descubrirá, es que la ‘teoría de justicia’ de Rawls no funciona y que muchos de sus argumentos son endeble.”<sup>73</sup>

Por otro lado, pienso que la crítica de Barry también está dirigida al liberalismo en general. Sin embargo, el autor concluye expresando su fuerte apego al liberalismo en cuanto ideas, indicando que comulga con los valores de libertad individual, responsabilidad y autonomía. El tema, me parece, es que aun cuando Barry profesa valores liberales, no está de acuerdo en que estos mismos valores sean utilizados

---

<sup>73</sup> *Ibidem*, p. ix

como premisas para construir una teoría que, desde su particular punto de vista, propone como resultado dos principios, a partir de los cuales estas mismas premisas pueden ser derivadas.

Sería muy complejo entrar en profundidad a revisar cada uno de los argumentos de Barry, que en estricta lógica me parecen correctos, en contra de la teoría rawlsiana. De hecho, en la sección 18 hice notar algunos ejemplos de cómo Rawls introduce en varias instancias restricciones en la posición original que finalmente favorecen la elección de sus principios. Sin embargo, pienso que su enfoque de elección racional estricto resulta demasiado estrecho para valorar el contenido de una propuesta tan compleja que involucra no sólo lo racional, sino también lo moral, puesto que lo ético no puede ser reducido a una argumentación sistemática racional. Argumento que Rawls está equivocado cuando pretende enmarcar su propuesta dentro de la teoría de decisión racional al expresar: “La teoría de la justicia es una parte, quizá la más significativa, de la teoría de la elección racional.”<sup>74</sup> La ética y la justicia no pueden ser parte de una teoría estricta racional, pues los conflictos morales inherentes a la existencia del hombre, con sus rasgos contradictorios, no pueden ser siempre decididos por la lógica contundente. Así, la crítica de Barry es provocada por el propio Rawls al pretender dar a su propuesta una validez racional que no tiene. En todo caso me parece que la verdadera contribución de la propuesta rawlsiana no estriba en su rigor racional, sino en su revitalización del contrato social como una herramienta que permite desarrollar y llevar el concepto de la justicia a nuevas fronteras, y en su crítica a las deficiencias del utilitarismo.

### Crítica de Habermas

En su ensayo *Reconciliación Mediante el Uso Público de la Razón*<sup>75</sup>, Habermas emprende una aguda crítica a la teoría de justicia rawlsiana, dentro de un debate que el filósofo alemán describe como una ‘disputa en familia’, dado que ambos autores son pensadores reconocidos de la línea liberal.

---

<sup>74</sup> *Op. cit.*, p.29

<sup>75</sup> Jürgen Habermas, “Reconciliación Mediante el Uso Público de la Razón” en *Debate Sobre el Liberalismo Político*, Tr. Gerard Vilar Roca. Barcelona: Paidós, 1998.

Procedo a hacer la revisión de la crítica de Habermas, en tres partes. La primera objeción se dirige contra el diseño de la posición original, en tres flancos. Para Habermas la teoría de la elección racional desde la cual se construye la posición original, no es un planteamiento adecuado para dar cuenta de la racionalidad práctica, y más específicamente, de la 'justa reciprocidad' de las obligaciones normativas. Un segundo elemento de objeción es la perspectiva individualista de la posición original que propone reflexionar sobre lo normativo reduciéndolo a intereses y bienes. Esto tiene como resultado que los derechos sean incluidos por Rawls dentro de los bienes primarios a distribuir, contraviniendo su esencia de ser disfrutados sólo en la medida en que se ejercen. Para Habermas, los derechos regulan relaciones entre agentes, y no pueden ser 'poseídos' como cosas. La tercera parte de esta primera impugnación se lanza sobre el artificio del velo de ignorancia, que priva a las partes representantes de la posibilidad esencial de una perspectiva intersubjetiva, que permitiría comprender la posición de los otros, para llegar a un acuerdo justo. Al respecto, comenta Habermas:

Tras la fundamentación de los principios de justicia, el velo de ignorancia es progresivamente levantado en los sucesivos pasos de las constituyentes, de la legislación básica y de la aplicación del derecho. Puesto que entonces las informaciones que fluyen en torrente tienen que armonizar con los principios ya elegidos en condiciones de privación de información, no debe haber ninguna sorpresa desagradable. Si queremos asegurarnos frente a ello, tenemos que construir la posición original ya con conocimiento e incluso con previsión, de todos los contenidos normativos que pudieran representar en el futuro un potencial que estimule la comprensión compartida de sí mismos y del mundo de ciudadanos libres e iguales.<sup>76</sup>

Así, la justificación normativa rawlsiana asume el conocimiento del mundo en el diseño, previo a la elección de los principios de justicia, de una sociedad bien ordenada, que brinda a los miembros de la sociedad un todo ya resuelto, vulnerando así su autonomía cognitiva al formular una teoría de la justicia que pretende ser sustantiva. Habermas, en contraste, a través de la ética discursiva busca integrar el punto de vista moral de la sociedad en el proceso democrático de una argumentación intersubjetiva, entre sujetos iguales, pero con convicciones morales diferentes, bajo la premisa de un discurso

---

<sup>76</sup> *Ibidem*, p.53

inclusivo y libre de coerción, en el que cada una de los miembros de la sociedad propondrá su punto de vista. Así, el diseño de la sociedad y los principios de justicia surgen del discurso, y no vienen dados como algo ya establecido.

La segunda objeción busca exhibir una falta de claridad en los cimientos teóricos del giro político dado por Rawls a su teoría de justicia, en el que modifica su propuesta inicial para delimitarla a lo político, eliminando el alcance metafísico que originalmente pretendía. Abre la primera sección atacando los dos conceptos que sustentan este giro político: el consenso por superposición y la razón pública. Para Habermas, la definición del consenso en dos etapas, una constitucional y otra legislativa, presupone una teoría del orden político anterior a la participación democrática, lo que anularía sus pretensiones de legitimación al imponer un precepto objetivo por encima del principio de soberanía popular democrática. La segunda parte critica la visión rawlsiana de 'lo razonable', que es presentado como elemento de la construcción política de una razón práctica frente a una razón teórica. La dificultad con esta definición de la racionalidad práctica como 'lo razonable', es su proyección exclusivamente política y su pretensión de validez pública. Por lo tanto, si la validez no está ya respaldada por elementos cognitivos ni morales, sino únicamente por lo político, lo público de la teoría normativa debe lograr su legitimación en la premisa virtuosa de una sociedad bien ordenada.

La tercera y última discrepancia gira alrededor de la disputa entre la 'libertad de los antiguos' –entendida como los derechos políticos republicanos de participación que hacen posible la autodeterminación de los ciudadanos- y la 'libertad de los modernos' – el núcleo de los derechos subjetivos privados que incluye la libertad de creencias y conciencia, la protección de la vida, la libertad personal y la propiedad. Para Habermas, la presentación de la teoría rawlsiana en dos etapas genera una prioridad a favor de los derechos liberales que subordina el proceso democrático indispensable a una categoría inferior, lo cual es inaceptable. Por otro lado, la autonomía modelada en la posición original queda limitada a una versión puramente virtual, pues no puede ser desplegada en una sociedad constituida jurídicamente, dado que en la medida que se levante el velo de ignorancia y las partes de la posición original se conviertan en personas de carne y hueso, los miembros de la sociedad rawlsiana se encontrarán más inmersos en una estructura progresivamente institucionalizada más allá de su control. En otras

palabras, desde la perspectiva de la justicia rawlsiana, la fundación del estado de derecho democrático no puede ser repetida bajo las circunstancias de una sociedad que ya ha sido previamente bien ordenada, en la que el uso de la razón pública no juega un papel relevante en el ejercicio de la autonomía política, sino que está restringido a la misión de mantener la estabilidad política.

Estos son los elementos centrales de la crítica de Habermas. El filósofo estadounidense reacciona con la publicación de una réplica, que no cubriré en este espacio, a la que Habermas también da respuesta.

He seleccionado estas tres críticas, una desde el comunitarismo, otra desde la óptica de la teoría de la elección racional, y otra desde la perspectiva 'familiar' del liberalismo, con el afán de ilustrar la variedad de visiones críticas acerca de la propuesta de Rawls. En mi opinión, Rawls, al elaborar un edificio conceptual tan sofisticado y complejo, que buscó presentar una propuesta que respondiera a las objeciones que le fueron expuestas a lo largo de los más de veinte años de su elaboración, consiguió estructurar una obra monumental que ha tenido una gran influencia en los medios académicos de todo el mundo. Donde Rawls falla es en la pretensión de ofrecer una teoría racional estricta, como lo hace notar duramente Barry. Sin embargo, dado que la filosofía moral y la filosofía política no son ciencias duras, y que las diferentes corrientes de pensamiento filosófico postulan ópticas diversas, que no pueden ser armonizadas en un gran sistema, me parece hasta cierto punto natural el que haya fisuras en la teoría de justicia de Rawls, y que éstas hayan sido señaladas por pensadores incisivos como Sandel y Habermas. Hasta el día de hoy se siguen publicando numerosas críticas y comentarios a la propuesta de Rawls, lo que refleja la vigencia de sus reflexiones, su gran influencia, y el vigor de su pensamiento.

#### **24. LA TEORÍA RAWLSIANA EN UN CASO PRÁCTICO**

La propuesta rawlsiana es elegante y atractiva dada la claridad de sus dos principios. La pregunta natural que surge es si es posible hacer uso de la teoría de Rawls para resolver un caso de justicia social en la práctica, en el complejo terreno de lo político. Pienso que un caso interesante es el proceso de negociación de la reforma al sistema

de salud de Estados Unidos llevado a cabo por el presidente Obama, mejor conocida como Obamacare. El objetivo de esta reforma era mejorar la situación de cuarenta y seis millones de personas que en 2008 no tenían seguro médico; en términos rawlsianos, el plan del presidente era proponer una legislación que resultara en una mejora para los menos favorecidos del sistema.

Está documentado que Obama, quien estudió leyes en la Universidad de Harvard, fue alumno de Rawls y se vio fuertemente influenciado por sus ideas. Algunos conceptos rawlsianos pueden ser identificados en el discurso que dio el mandatario al abrir la Cumbre de Salud de 2009 en la que se iniciaron las negociaciones políticas. El presidente dijo que había venido a Washington a trabajar por la gente, por encima y en contra de 'esos grandes intereses'. La implicación de esta frase es clara: para que se logre la justicia, se debe promover un resultado justo para la gente, y en contra de los intereses de los grupos poderosos.

Pienso que otros rasgos rawlsianos son visibles cuando Obama intenta establecer guías generales para acuerdos que no son específicos: todas las voces serán escuchadas, y todas las ideas y opiniones serán bien recibidas; se debe aceptar la necesidad de llegar a compromisos, pues no es factible que todos los interesados – empresas, partidos políticos, gobernante y representantes de los ciudadanos- obtengan lo que quieren; y nadie debe dominar las discusiones, que deben ser incluyentes. Asimismo, no se permitirá la continuación de la situación de injusticia que prevalecía en ese momento.

La aplicación del pensamiento rawlsiano se empieza a dificultar cuando los temas se vuelven más específicos –cómo hacer los cambios de legislación, y cuándo hacerlos. En el caso de la reforma de salud, al irse dando a conocer los detalles más específicos sobre la propuesta de Obama, surgieron muchas preguntas e ideas alternativas por parte de los republicanos, que requerían ser respondidas y debatidas. Estas inquietudes no fueron seriamente consideradas y el consenso inicial de partidos se fracturó. Planteo que la interpretación rawlsiana de Obama consideró inicialmente que las visiones diferenciadas de las dos fuerzas políticas habían llegado a un acuerdo que formaba parte del consenso traslapado. Sin embargo, este convenio exige que las partes presenten posturas 'razonables'. En el caso de la reforma de salud, interpreto

que Obama juzgó como 'no razonable' la postura republicana conocida como 'empezar de nuevo', que permitiría mantener el status quo por más tiempo. Así, los republicanos no fueron merecedores de ser escuchados. De hecho, como ya mencioné, en sus palabras de apertura Obama advierte que no permitiría tres situaciones: que alguna postura domine las discusiones; la oposición al cambio; y mantener el status quo. Argumento que el presidente estaba delimitando la frontera entre lo razonable y lo no razonable del consenso traslapado.

La postura radical de Obama, que ignoró las demandas republicanas produjo un rechazo total del partido opositor, y una polarización política en Estados Unidos que se vive hasta el día de hoy. Respecto a la reforma, después de muchos y azarosos esfuerzos por parte de los demócratas, ésta fue aprobada por el senado con la totalidad de los cincuenta y ocho votos democráticos, el apoyo de dos senadores independientes, y el rechazo de los cuarenta representantes republicanos.

El costo para Obama de intentar poner en práctica una filosofía académica que no está anclada en la realidad, a una situación compleja que requería del proceso negociador político, fue alto y se refleja en la caída de su popularidad y el deterioro de su imagen. Adicionalmente, la aprobación polarizada le ha causado a Obama un rechazo permanente por parte de los republicanos que le ha dificultado la gestión de su gobierno. Esto es evidente en los intentos anuales que el partido conservador lleva a cabo para derogar el Obamacare, que ha quedado marcado como un ejemplo invertido de la política, entendida como el arte de lo posible.

## **25. EL PRINCIPIO DE DIFERENCIA Y SU APLICACIÓN**

En este apartado hago un acercamiento incremental a mi tema central, la crítica de Harsanyi, al exponer lo relacionado con el principio de diferencia, que incluye en su redacción un uso limitado de la regla maximin, relacionado con los menos favorecidos.

El núcleo del principio de diferencia es algo simple y atractivo: las desigualdades sociales y económicas deberían ser evaluadas en términos de qué tan bien quedan los que están hoy peor. La idea es bastante sencilla, pues solamente pide que el valor mínimo del algún índice, que representa a los menos favorecidos, sea maximizado. Para muchos este principio es muy atractivo, pues requiere que las ventajas disfrutadas

por los que están en el nivel más bajo, sean tan generosas como sea posible, articulando de una manera transparente y elegante la preocupación principal del pensamiento igualitario. Tiene además la ventaja de evitar dos extremos que pueden caer en el absurdo: la igualdad a cualquier precio, que ha demostrado ser imposible de lograr, o la maximización del agregado total, sin importar nada cómo sea distribuido, dejando posiblemente a los más desaventajados en condiciones de pobreza extrema que no les permitiría vivir con plenitud razonable su condición de seres humanos. Rawls ofrece un esquema intermedio: pueden tener más bienes los talentosos y los más ambiciosos, y la distribución de esos bienes puede ser muy desigual, siempre y cuando los menos favorecidos estén mejor.

Recordemos los dos principios de justicia de Rawls, ya enunciados en su forma definitiva:

#### *Primer Principio*

Cada persona ha de tener un derecho igual al más extenso sistema total de libertades básicas compatible con un sistema similar de libertad para todos.

#### *Segundo Principio*

Las desigualdades económicas y sociales han de ser estructuradas de manera que sean para:

- a) mayor beneficio de los menos aventajados, de acuerdo con un principio de ahorro justo, y
- b) unidos a los cargos y las funciones asequibles a todos, en condiciones de justa igualdad de oportunidades.

Estos principios se aplican, en primer lugar, como ya se ha mencionado, a la estructura básica social, y ordenan la asignación de derechos y obligaciones administrando la distribución de las ventajas económicas y sociales. La formulación de estos dos principios presupone que en la sociedad existen dos esferas diferentes: las libertades, y los aspectos que establecen desigualdades económicas y sociales; el

primer principio refiere a la primera esfera, mientras que el segundo se conecta con la segunda.

Es importante notar que estos dos principios se aplican en un orden lexicográfico que establece que las libertades del primer principio tienen prioridad sobre la distribución referida en el segundo. En otras palabras, no es aceptable para la justicia rawlsiana una restricción de libertades a cambio de una porción mayor de bienes materiales.

En particular, respecto del segundo principio, Rawls escribe:

El segundo principio se aplica...a la distribución del ingreso y la riqueza y a formar organizaciones que hagan uso de las diferencias de autoridad y responsabilidad o cadenas de mando. Mientras que la distribución del ingreso y de las riquezas no necesita ser igual, tiene no obstante que ser ventajosa para todos, y al mismo tiempo los puestos de autoridad y mando tienen que ser accesibles a todos. El segundo principio se aplica haciendo asequibles los puestos y, teniendo en cuenta esta restricción, disponiendo las desigualdades económicas y sociales de modo tal que todos se beneficien.<sup>77</sup>

Así, este segundo principio exige regular no solamente los elementos relacionados con el bienestar material, sino también aquellos concernientes al acceso a puestos de autoridad y mando, dado que desde esas posiciones los miembros de la sociedad pueden lograr condiciones materiales o de poder ventajosas, que generen circunstancias injustas entre los ciudadanos.

La formulación precisa del principio de diferencia corresponde a la primera parte del segundo principio, que dice 'las desigualdades económicas y sociales han de ser estructuradas de manera que sean para mayor beneficio de los menos aventajados, de acuerdo con un principio de ahorro justo.'

El principio de *ahorro justo* es un mecanismo teórico que establece Rawls con el objeto de que las perspectivas de bienestar de los menos favorecidos se puedan extender a las generaciones futuras. El argumento es que cada generación no solamente debe conservar las ventajas de la vida cultural y mantener en excelente estado las instituciones justas establecidas, sino además realizar en cada período una

---

<sup>77</sup> *Op. cit.* p.68

cantidad significativa de acumulación de capital, a través del ahorro. Este ahorro puede ser utilizado de varias maneras, en maquinaria, en otros medios de producción, y hasta en servicios fundamentales para el desarrollo social, como la educación. Así, el principio de *ahorro justo* nos indicaría qué amplitud tendrá la inversión requerida para determinar un nivel social mínimo, y por tanto qué cantidad de capital será necesario y justo ahorrar anualmente. De ahí que el principio de diferencia tenga este rasgo cualitativo.

## **26. LA REGLA MAXIMIN EN EL PRINCIPIO DE DIFERENCIA**

El principio maximin surge en el terreno de la teoría de juegos, rama de las matemáticas aplicadas que se encarga de utilizar modelos para estudiar las interacciones de los jugadores racionales, que se asume son seres humanos, en condiciones de estructuras formalizadas de incentivos. El principio maximin es una estrategia dentro de la teoría de juegos, cuya recomendación es maximizar el valor mínimo. En este sentido el principio de diferencia puede ser interpretado como una aplicación de la regla maximin, pues su objetivo es precisamente maximizar el valor mínimo de bienestar de los menos favorecidos, en el contexto de una sociedad que diseña, desde la posición original, los principios de justicia que han de regirla.

Es importante diferenciar esta aplicación específica del criterio maximin dentro del principio de diferencia, de la aplicación de maximin como criterio de elección de principios de justicia, a utilizarse por las partes en su proceso de evaluación de alternativas. Recordemos que la crítica de Harsanyi está enfocada precisamente a su uso en este sentido amplio.

## **CAPÍTULO 2. LA CRÍTICA DE HARSANYI**

En esta parte haré una descripción de la crítica de Harsanyi. También presentaré los puntos de vista de varios comentaristas que han escrito sobre el tema. En el último capítulo de conclusiones haré mi análisis personal e intentaré llegar a mis propias conclusiones.

La reprobación de Harsanyi surge desde la perspectiva del ámbito de la teoría de elección racional, un campo que no está normalmente vinculado a la filosofía. Por tanto, me parece relevante antes de adentrarme en la crítica, responder a la pregunta: ¿qué papel juega la teoría de elección racional dentro de la propuesta rawlsiana? La respuesta la encuentro en la postura que el mismo Rawls adopta cuando afirma que la teoría de la justicia es una parte importante de la teoría de la elección racional. Resulta pues evidente que Rawls considera su concepción de justicia como insertada dentro del ámbito de la elección racional, para enfatizar que los individuos de su sociedad son agentes racionales morales. Como ya mencioné, soy de la opinión que esta pretensión de Rawls es equivocada.

## **27. LA INCERTIDUMBRE: PUNTO DE PARTIDA DE HARSANYI**

El punto de partida de la crítica del Nobel es la posición original rawlsiana, en la que las partes representantes, debido al velo de ignorancia, no conocen su posición social y económica, ni sus intereses en la sociedad, ni tampoco sus talentos personales y habilidades. Desde esta posición de incertidumbre, las partes deben elegir los principios de justicia que regirán a la sociedad. Así, para Rawls, el velo de ignorancia conduce directamente al problema de la elección en incertidumbre total.

La escuela utilitarista establece que la regla de decisión a usar por individuos racionales en una situación de incertidumbre es la maximización de la utilidad esperada, con una distribución bayesiana<sup>78</sup> de probabilidades. Rawls, por el contrario, argumenta que el grado de incertidumbre es tan elevado que no es posible ni siquiera asignar alguna probabilidad a los posibles escenarios alternativos que tendría la sociedad, ni a los posibles lugares que ocuparían los individuos representantes que deben elegir principios de justicia. En efecto, dice Rawls:

---

<sup>78</sup> Bayesiana refiere a que se asigna una probabilidad similar a cada una de las posibles situaciones de arreglo social, y también a cada una de las posibles posiciones del individuo representante de los ciudadanos, dentro de la sociedad a establecer.

el velo de ignorancia excluye todo conocimiento acerca de las probabilidades. Las partes no tienen ninguna base para determinar la naturaleza probable de su sociedad, ni de su lugar en ella. No tienen, por tanto, bases para el cálculo de probabilidades.<sup>79</sup>

Esta imposibilidad total de asignar probabilidades en la posición original, lleva a Rawls a proponer el controvertido uso de la regla maximin, cuyo uso como criterio de elección de principios de justicia es severamente cuestionado por Harsanyi.

## **28. ELEMENTOS DE LA CRÍTICA DE HARSANYI**

Harsanyi establece desde el inicio de su crítica a maximin que su objetivo es argumentar que “el intento de Rawls de sugerir una alternativa viable al utilitarismo no tiene éxito”<sup>80</sup>; es decir, para Harsanyi el uso de la regla maximin, por ser irracional, hace que toda la propuesta rawlsiana se desmorone.

A partir de un ejemplo específico en el que una persona debe escoger entre dos trabajos, uno de bajos ingresos en Nueva York, donde vive el individuo, y otro en Chicago con mejor remuneración, a donde tendría que viajar en avión -con el consiguiente riesgo de accidente mortal-, Harsanyi busca demostrar que el uso de la regla maximin lleva a decisiones irracionales. Es claro que el viaje a Chicago tiene riesgo de un accidente aéreo, aun cuando las probabilidades de que este ocurra sean muy bajas. El argumento de Harsanyi se centra en el hecho de que, al no asignar probabilidades al evento de morir en el vuelo a Chicago, el individuo asumirá que sus opciones son: a) trabajo en Nueva York con bajo salario y b) trabajo en Chicago con mejor salario, pero con muerte segura. Por tanto, siguiendo el criterio maximin, la persona considerará que es menos mala la opción de tomar el trabajo en su ciudad. Por otro lado, si se aplicaran las probabilidades correspondientes, y el criterio utilitarista de maximizar el valor esperado -definido como la utilidad de cada alternativa, multiplicado por la probabilidad de que ella ocurra-, la mejor decisión será tomar el vuelo a Chicago y obtener el trabajo más benéfico, dado que la probabilidad estadística

---

<sup>79</sup> *Op. cit.*, p.152

<sup>80</sup> *Op. cit.*, p. 594.

de que suceda un accidente aéreo en el viaje a Chicago es extremadamente baja. Esta conclusión concuerda con el punto de vista práctico intuitivo.

Otro ejemplo es presentado para mostrar la debilidad moral del principio de diferencia, que lleva a decisiones inaceptables. Existe una sociedad con tres individuos, un médico y dos pacientes que están gravemente enfermos de neumonía. La única posibilidad de recuperarse es con un tratamiento a base de antibióticos, pero sólo hay una dosis disponible para tratar a un enfermo. Uno de los dos pacientes, el individuo A, es una persona sana, quitando el ataque actual de neumonía, mientras que el individuo B sufre de un cáncer terminal, pero podría prolongar su vida algunos meses con el tratamiento del antibiótico. Harsanyi pregunta ¿a qué persona se debería dar la medicina? Si se sigue el principio de diferencia, que da prioridad a los menos afortunados, el antibiótico debería ser brindado al individuo B, que está en una situación más grave. En contraste, el utilitarismo, que busca la maximización del beneficio, propondrá que se salve al paciente A, pues esto producirá un mayor provecho social al regresarlo a una vida plena, puesto que el individuo B morirá de cualquier manera en poco tiempo. La decisión sugerida por el principio maximin es, para Harsanyi, una decisión moralmente inaceptable. Nuevamente, el economista apela al sentido común, con lo cual concuerdo.

A partir de estos dos ejemplos y otro parecido las principales conclusiones de Harsanyi son:

- a) Dado que la regla maximin lleva a decisiones irracionales, ésta debe ser rechazada como regla de decisión para la elección de principios de justicia para la sociedad, situados en la posición original;
- b) “El principio de diferencia con frecuencia tiene implicaciones morales completamente inaceptables”<sup>81</sup>;
- c) Las dificultades resultantes del uso de la regla maximin pueden ser evitadas si la regla de decisión es la maximización de la utilidad esperada de acuerdo al principio de la teoría Bayesiana (de iguales probabilidades);

---

<sup>81</sup> *Ibidem*, p.598

- d) Por tanto, el utilitarismo es una mejor propuesta de justicia que los dos principios de Rawls, y la teoría de justicia de Rawls debe ser descartada.

El texto de Harsanyi cubre algunos otros temas laterales, incluyendo una refutación de los argumentos de Rawls en contra del uso de cualquier noción de probabilidades en la situación inicial, y su razonamiento para descalificar al utilitarismo. No entraré en más detalle puesto que el núcleo de esta investigación es la descalificación del uso del principio propuesto por el pensador americano.

## 29. RESPUESTA DE RAWLS

Un año después de la publicación de la crítica, Rawls responde<sup>82</sup> con un breve texto de cinco páginas. La respuesta de Rawls inicia lanzando un ataque directo al método utilizado por Harsanyi, que presenta ejemplos específicos en los que no funciona la regla maximin, cuando dice: “el criterio maximin no está hecho para ser aplicado a situaciones de pequeña escala, como el caso de cómo un doctor debiera tratar a sus pacientes... maximin es un criterio macro, no micro”<sup>83</sup>. Pienso que Rawls evade dar una respuesta puntual en el terreno racional a las objeciones de Harsanyi, puesto que sabe que su teoría puede llevar a respuestas equivocadas en situaciones particulares y que, en las circunstancias de los ejemplos presentados por el economista, no es posible rebatir las conclusiones reprobatorias.

Rawls prosigue llevando la argumentación al terreno macro, proponiendo dos comparaciones consecutivas, cada una entre dos alternativas. La primera, reiterando la propuesta de su obra *Teoría de la Justicia*, confronta sus dos principios de justicia contra el utilitarismo que maximiza la utilidad promedio. Haciendo un breve repaso de los argumentos de su obra concluye que los dos principios son una propuesta superior al utilitarismo. Hasta aquí no hay nada nuevo.

---

<sup>82</sup> John Rawls, “Concepts of Distributional Equity: Some Reasons for the Maximin Criterion”. *The American Economic Review*, Vol. 64, No. 2, Papers and Proceedings of the Eighty-sixth Annual Meeting of the American Economic Association, May, 1974, p. 141-146.

<sup>83</sup> *Ibidem*, p.142.

Continúa Rawls planteando una segunda comparación muy novedosa: por un lado los dos principios de justicia rawlsianos y, por el otro, una reformulación de ellos en la que el segundo principio se sustituye por lo siguiente: “las desigualdades sociales y económicas deben ser modificadas de manera que se maximice la utilidad promedio de una forma consistente con una equitativa igualdad de oportunidad.”<sup>84</sup> Rawls argumenta a favor de los dos principios originales de justicia como se publicaron en 1971, a partir de la premisa de que en la posición original las partes serán sumamente sensibles al riesgo, dada la naturaleza tan crucial de su situación, pues una vez que sean elegidos los principios de justicia, estos tendrán vigencia permanente y no podrán ser cambiados por otros. Por otro lado, se apoya en un hallazgo de Arrow, el cual postula que “el estándar de la utilidad se acerca a maximin cuando la aversión al riesgo aumenta sin límites”.<sup>85</sup> Aunque Rawls no lo dice explícitamente, la aparente conclusión que el lector infiere es que de alguna manera las dos formulaciones de esta segunda comparación parecen equivalentes. Sin embargo, dice Rawls, “asumo que, basado solamente en consideraciones de aversión al riesgo, hay una diferencia significativa entre las dos concepciones. Por tanto, el problema es identificar otras características atractivas del criterio maximin que inclinen la balanza a su favor.”<sup>86</sup> Así, Rawls presenta adicionalmente cuatro argumentos de tipo práctico para defender al criterio maximin, dado que en situaciones de aversión moderada al riesgo no es clara la superioridad de esta regla.

El primero es que se requiere mucha menos información para aplicar el criterio maximin, pues la identificación de los menos afortunados puede hacerse relativamente fácil. Asumo que Rawls propondría alguna métrica de ingreso económico para ello, pues ya lo ha planteado en otras partes. Para Rawls resulta muy complejo definir qué es lo que maximiza la utilidad, pues se requiere una comparación precisa de las utilidades de los diferentes grupos sociales por medio de algún estándar relevante.

---

<sup>84</sup> *Ibidem*, p.143

<sup>85</sup> Ken Arrow, Rawls's Principle of Just Saving, en Institute for Mathematical Studies in the Social Sciences, *Stanford University, Technical Report no. 106*, Sep. 1973a.

<sup>86</sup> *Op. cit.*, p.143.

Este mecanismo dejaría tantos elementos sujetos a juicio que aquellos que sean favorecidos por las desigualdades actuales tendrían oportunidad de explotar sus ventajas y así aumentar aún más la injusticia de la sociedad.

Un segundo argumento refiere a la necesidad de que el criterio de distribución deberá de ser público, y por tanto comprensible para los ciudadanos promedio. De los posibles criterios que tengan claridad pública, el de la igualdad total es el que tiene mayor agudeza. Sin embargo, la igualdad completa no es logable. Para Rawls, el utilitarismo no tiene suficiente 'agudeza', pues aun cuando fuera logable maximizar la utilidad social total, habría poca confianza pública en su validez, dada la complejidad para demostrar su efectividad al ciudadano no conocedor. El principio maximin, en contraste, es suficientemente agudo para tener aceptación pública.

Otra justificación a favor de maximin es la condición de que los principios deben ser tolerables de manera permanente para todos los miembros de la sociedad. Para el caso de los menos aventajados, el principio de utilidad les pide que vean como una razón de que ellos tengan menos posibilidades de progresar, el que los otros tengan más ventajas y beneficios. Esto es extremadamente demandante desde el punto de vista psicológico. Por el contrario, el criterio maximin asegura a los menos favorecidos que las desigualdades jugarán a su favor. Así, el principio maximin permite que sea más factible que todos los miembros de la sociedad estén contentos con su situación.

Finalmente, Rawls plantea un cuarto argumento, después de aceptar que sus justificaciones anteriores no son decisivas de una manera contundente. Así, plantea como convincente la mayor capacidad de maximin para ayudar a alcanzar las aspiraciones de los ciudadanos para poseer una personalidad libre e igual. Dado que las perspectivas de progreso de vida en la sociedad están muy influenciadas por factores contingentes como el origen familiar y de clase social, y los talentos naturales, debemos preguntarnos bajo qué principio de justicia podrán las personas libres e iguales permitir que sus relaciones se vean afectadas por circunstancias de suerte social y lotería natural. Las personas desearán que los efectos de la suerte sean regulados por algún principio razonable. Bajo la regla maximin, postula Rawls, la distribución natural de talentos podría ser vista, en algunos aspectos, como un bien colectivo, pues la atención que se pondría en los menos favorecidos operaría como un

estabilizador de las consecuencias desfavorables para aquellos que no nacieron con talentos elevados, o en sectores sociales donde la educación de calidad no está disponible. Rawls admite que esta aplicación de maximin tendría enormes dificultades en la práctica, pues podría ser imposible medir la habilidad con alguna métrica concreta, y existe el riesgo de que las personas tuvieran incentivos para no revelar sus talentos.

En conclusión, Rawls plantea que maximin es un “punto focal natural entre la igualdad estricta y el principio de la utilidad promedio”<sup>87</sup>, que tiene muchas características atractivas. Sin embargo, a punto seguido comenta, en lo que me parece una manifestación de cierta duda sobre la solidez lógica de su argumentación: “no deseo sobre enfatizar este criterio (maximin): una investigación más profunda que cubra otras comparaciones entre dos alternativas tal vez muestre que hay alguna otra concepción de justicia más razonable.”<sup>88</sup> Más adelante daré mi punto de vista sobre la respuesta que da Rawls a Harsanyi.

Para cerrar la cronología del debate, me parece relevante considerar los comentarios de Harsanyi a la réplica de Rawls. En efecto, el economista es severo cuando evalúa el argumento de Rawls en el sentido de que maximin no es un principio para ser aplicado en situaciones de pequeña escala, y escribe: “Desgraciadamente, debo decir que ésta es una defensa particularmente inepta...Estoy realmente asombrado de que un filósofo distinguido como Rawls pudiera dejar de ver el simple hecho de que los contra ejemplos que he presentado (y los muchos otros ejemplos que uno pudiera fácilmente proponer) no tienen nada que ver con la escala.”<sup>89</sup> Harsanyi lleva la argumentación a su zona de competencia, las matemáticas, y cuestiona si para Rawls existe un valor  $x$ , a partir del cual las circunstancias de la justicia dejan de ser de pequeña escala, y por tanto de tamaño suficiente para merecer la aplicación del principio maximin. Y si se propone ese valor  $x$ , ¿cuáles serían las consideraciones

---

<sup>87</sup> *Ibidem*, p. 145.

<sup>88</sup> *Idem*

<sup>89</sup> John Harsanyi, *Essays on Ethics, Social Behavior, and Scientific Explanation*. D. Reidel Publishing Company. Holland: 1980, p. 59.

morales para determinar este número frontera? Así, dejando la pregunta abierta, Harsanyi concluye que la teoría de Rawls queda completamente descalificada como un competidor serio de la teoría utilitarista.

Después de la descalificación final por parte de Harsanyi, el debate ha sido objeto de estudio de muchos comentaristas; unos han dado la razón al filósofo y otros se la han dado al economista. En la siguiente sección revisaré algunas de estas posturas, para más adelante plantear mi propia valoración.

### **30. COMENTARISTAS SOBRE LA DISPUTA RAWLS-HARSANYI**

En este apartado haré una revisión de los puntos de vista presentados por dos comentaristas: por un lado, la visión de la profesora Kristine Schrader-Frechette, catedrática de la Universidad de Notre Dame, quien publicó una argumentación extensa<sup>90</sup> del debate desde la perspectiva de la filosofía y de la teoría de decisión racional; por el otro, la postura del académico británico Ken Binmore, reconocido teórico de la decisión racional. Mi objetivo es ampliar la perspectiva de la disputa con estas ópticas adicionales.

Schrader-Frechette

La profesora norteamericana plantea en su obra, publicada en 1991, que “las reglas prevalecientes Bayesianas utilitaristas están con frecuencia equivocadas”; y, agrega la autora, “existen razones convincentes para rechazar la estrategia Bayesiana o utilitarista bajo incertidumbre, y es con frecuencia más racional preferir la estrategia maximin.”<sup>91</sup> Más adelante emprende el análisis de los argumentos de los protagonistas del debate.

A partir del ejemplo ofrecido por Harsanyi en su crítica, relacionado con la decisión entre dos trabajos, uno en Nueva York, y otro en Chicago, la autora despliega una lógica sistemática que conjuga con ejemplos en los que la regla maximin lleva a conclusiones racionales, para refutar los argumentos del economista y sus

---

<sup>90</sup> Kristine Schrader-Frechette, *Risk and Rationality, Philosophical Foundations for Populist Reforms*, Los Angeles. University of California Press, 1991.

<sup>91</sup> *Ibidem*, p.102

conclusiones. La premisa clave del nobel, y sostén del modelo utilitarista, es también objetada: “la premisa de equiprobabilidad es también éticamente cuestionable puesto que no permite el asignar un peso igual a priori a los intereses de todos los individuos, como Harsanyi pretende.”<sup>92</sup> Es importante notar que el punto esencial de conflicto entre Rawls y el economista es el uso de las probabilidades en sus modelos. Para Rawls la situación inicial es de una incertidumbre tan elevada que no es posible asignar ningún tipo de probabilidades a las posibles configuraciones sociales, mientras que para Harsanyi, la incertidumbre total implicaría la asignación de probabilidades idénticas a cada una de las posibles circunstancias sociales. Para Schrader-Frechette, solamente en el caso en que “en realidad se tiene información parcial, entonces el uso de la premisa de equiprobabilidad, junto con la estrategia Bayesiana/utilitarista, hace más sentido”<sup>93</sup> que la regla maximin.

El desarrollo de la evaluación de los méritos de los argumentos de Harsanyi y Rawls es bastante elaborado, y en algunos partes hace uso de elementos técnicos de la teoría de riesgo que no voy a mencionar por estar fuera del alcance de este trabajo. Lo relevante es que, a los ojos de la filósofa americana, los postulados acerca de la irracionalidad de la regla maximin, y las pretendidas consecuencias morales inaceptables de ese criterio no son del todo válidos, pues se pueden refutar desde varios frentes.

Respecto a los argumentos de Rawls a favor de la regla maximin, estos son resumizados como (1) permite dar la mayor prioridad a los intereses de los menos aventajados; (2) evita usar la noción de función de utilidad, diseñada para el campo del análisis de riesgos y economía, en el área de lo moral, donde no pertenece; y (3) evita el uso de las comparaciones interpersonales Bayesianas/utilitaristas en la definición de justicia. Después de una rigurosa discusión, Schrader-Frechette concluye a favor de maximin cuando dice:

Estos argumentos acerca de los méritos de maximin y de las estrategias Bayesianas/utilitarias para decisiones bajo incertidumbre sugieren varias razones por las que el principio maximin podría ser

---

<sup>92</sup> *Ibidem*, p.104

<sup>93</sup> *Ibidem*, p.105

superior en ciertas situaciones. Por un lado, el Bayesianismo/utilitarismo podría exigir demasiado o muy poco de los agentes. Exigiría demasiado al comprometerlos a actos supererogatorios, como lo requiere la maximización de la utilidad. Podría exigir muy poco de los agentes al comprometerlos a maximizar la utilidad promedio, incluso en el caso de una tiranía de una mayoría contra una minoría. Adicionalmente, aunque la posición maximin es susceptible de sus propias dificultades (por ejemplo, asume que podemos especificar preferencia ordinal (de los principios), esos problemas con frecuencia parecen menos dificultosos, especialmente en situaciones de incertidumbre, que aquellos inherentes al enfoque Bayesiano.<sup>94</sup>

La pensadora americana termina analizando algunas consideraciones de tipo práctico relacionadas con la normatividad estadounidense sobre los temas ambientales, la tendencia de las autoridades gubernamentales a actuar de manera lenta e ineficiente en casos de emergencia, causando así daño a los sectores de la población más expuestos, y la frecuente falla de los más calificados expertos al diseñar sistemas que contienen elementos riesgosos. Su conclusión final es: “Por definición, no podemos evitar errores y corregir nuestros modelos en situación de incertidumbre. Por lo tanto, la prudencia prescribe que utilicemos reglas conservadoras de evaluación de riesgo, como maximin.”<sup>95</sup>

Así, Schrader-Frechette concluye dando mayor peso a los argumentos de Rawls, por encima de los de Harsanyi. Agrega la autora otras razones que en su perspectiva favorecen al uso del principio maximin en campos de riesgo e incertidumbre, más allá del campo teórico de la justicia, dando así peso a esta regla de decisión en el terreno práctico en el que es necesario ejercer la prudencia.

## Binmore

El académico británico Ken Binmore es un reconocido intelectual, matemático convertido en economista, y más adelante en teórico político, además de ser uno de los fundadores de la teoría de juegos. Este catedrático ha estudiado la disputa Rawls-Harsanyi, y ha escrito numerosas páginas al respecto. En una de sus obras más

---

<sup>94</sup> *Ibidem*, p.110.

<sup>95</sup> *Ibidem*, p.112.

conocidas<sup>96</sup>, Binmore deja claro desde un principio que está en desacuerdo con Rawls. En efecto, refiriéndose a la versión inicial de *TJ*, Binmore escribe: “Me gusta la descripción que Rawls hace acerca de la naturaleza de una sociedad justa, pero encuentro inaceptables los argumentos con los que defiende su postura.”<sup>97</sup>

Uno de los aspectos concretos de desacuerdo con Rawls es la manera en que *TJ* evita ofrecer una visión del estado de la naturaleza: cuáles son los rasgos del hombre, y qué consideraciones hay que tomar en cuenta sobre sus inclinaciones psicológicas y sociales para asumir que en la situación inicial se podrá llegar a un acuerdo sobre los principios de justicia. A los ojos de Binmore, “Rawls ve el instrumento de la posición original como un sustituto para el estado de la naturaleza, como lo entiende Hobbes.”<sup>98</sup>

Más adelante Binmore comenta negativamente sobre la argumentación de Rawls en contra del utilitarismo, en específico su propuesta de la regla maximin, resultado de la prohibición de cualquier tipo de probabilidad en el contexto de la posición original, y su planteamiento de utilizar un ‘índice de bienes primarios’ como sustituto de las curvas de preferencia del utilitarismo. Escribe Binmore:

Hay dos problemas con la revocación de Rawls de la teoría de utilidad...el primero es su uso del criterio maximin como un principio para los individuos que tomarán las decisiones bajo incertidumbre. El segundo es su idea sobre un índice de bienes primarios para hacerse cargo, entre otras cosas, de los problemas de las comparaciones interpersonales. En los dos temas, mi enfoque está basado en el de Harsanyi, cuya teoría de utilidad es completamente ortodoxa.<sup>99</sup>

Así, la conclusión del británico es claramente a favor de Harsanyi. El análisis de la crítica a maximin es bastante elaborada técnicamente, por lo que no entraré en detalles sobre ella. Lo que deseo resaltar es una frase que Binmore incluye en su texto, la cual

---

<sup>96</sup> Ken Binmore, *Game Theory and the Social Contract, Vol I: Playing Fair*. Boston: Massachusetts Institute of Technology, 1994.

<sup>97</sup> *Ibidem*, p.9.

<sup>98</sup> *Ibidem*, p.13

<sup>99</sup> *Ibidem*, p.51

ha sido citada en numerosos ensayos, que resume la visión del autor sobre la disputa: “Todos los académicos más influyentes, están del lado de Harsanyi. Su posición es completamente ortodoxa entre los teóricos de la decisión.”<sup>100</sup>

Me parece relevante resaltar que la evaluación de Binmore expresa dos cosas importantes y distintas: primero, que la mayoría de los expertos reconocidos en teoría de decisión racional han dado la razón a Harsanyi; sin duda existen algunos, además de Schrader-Frechete, que están del lado de Rawls, pero la mayoría de los comentaristas acreditados considera que la defensa de maximin es débil y que el ganador es Harsanyi. El segundo aspecto que deseo enfatizar, es que la perspectiva de Binmore, y la de los ‘académicos más influyentes’ es estrictamente desde el punto de vista de la aplicación de la teoría normativa de decisión. Es decir, que hasta el momento en que esta obra se publica, en el año de 1994, la óptica del debate estaba centrada en la metodología de la decisión, y no se había difundido otra manera diferente de estudiar el tema, en un terreno intelectual distinto. Más adelante volveré a este punto.

Es interesante notar que Binmore -aunque da la razón a Harsanyi en cuanto al principio maximin- no se considera a sí mismo miembro de la doctrina utilitarista; es más, él está en claro desacuerdo con el utilitarismo como un sistema integral de justicia. De hecho, a lo largo de la obra, el inglés camina en forma paralela, formulando sus críticas a Rawls, alternadas con frecuentes afirmaciones a favor del economista Nobel, y va construyendo una propia teoría de justicia, en la que busca adaptar los enfoques de Rawls y Harsanyi, y transformarlos en una nueva propuesta, una especie de fusión de las dos teorías en disputa. Dice Binmore:

En la raíz de mi propio enfoque constructivo hay una síntesis de las teorías de Rawls y se toma prestado el uso de Harsanyi de la teoría de la utilidad para incrustarla en un marco contractual como el concebido por Rawls.<sup>101</sup>

---

<sup>100</sup> *Ibidem*, p.292

<sup>101</sup> *Ibidem*, p.52

Binmore denomina a su nuevo sistema *Justicia Natural*, y publica<sup>102</sup> una versión más breve del mismo unos años más tarde. Esta propuesta busca combinar una concepción de lo moral influenciada por Hume, con la técnica de la teoría de juegos, y la teoría de la evolución. Esta parte de su texto ilustra la noción moral ‘científica’ propuesta por Binmore:

Las reglas morales que realmente gobiernan nuestro comportamiento consisten en una mezcla de instintos, costumbres y convencionalismos que son simultáneamente más mundanos y más complejos de lo que la escuela tradicional está dispuesta a aceptar. Ellas están moldeadas en su mayor parte por fuerzas evolucionarias -tanto sociales como biológicas. Si se desea estudiar estas reglas, no ayuda el preguntar cómo ellas coadyuvan a la causa de Dios, o a preservar lo Correcto. Lo que necesitamos es preguntar cómo evolucionaron esas reglas y por qué sobrevivieron. Es decir, necesitamos tratar a la moralidad como una ciencia.<sup>103</sup>

No profundizaré más sobre este tema por estar fuera del alcance de mi investigación, pero me parece ciertamente relevante que un académico de este calibre haya emprendido esta nueva teoría a partir de las propuestas confrontadas, buscando superar las debilidades que identificó en cada una de ellas.

En sinopsis, Schrader-Frechette está a favor de Rawls, pero Binmore apoya a Harsanyi, y deja claro que en el tema de la disputa, la mayoría de los expertos en teoría de decisión racional da la razón a Harsanyi. Hasta cierto punto, esto no es de extrañar, puesto que el utilitarismo del economista está fundamentado en la teoría de lo racional y hace uso de las herramientas de ésta para construir su sistema de justicia.

### **CAPITULO 3. CONCLUSIONES**

Procedo a exponer mis reflexiones finales. A partir de una revisión del papel que tienen lo moral y lo racional en la *TJ*, postulo que Rawls hace un esfuerzo extraordinario por

---

<sup>102</sup> Ken Binmore, *Natural Justice*. EUA: Oxford University Press, 2005.

<sup>103</sup> *Ibidem*, p.1

dotar a su propuesta de una racionalidad similar a la de su opositor, el utilitarismo. Propongo que Rawls comete un error en este sentido, que le causó numerosas críticas que tal vez podría haberse evitado. Más adelante argumento a favor de maximin como criterio de aceptación de una concepción mínima de justicia que cumpla con los condicionantes elevados de Rawls y presento mi conclusión sobre la disputa desde dos perspectivas. Primeramente, desde la óptica estrictamente racional, me declaro a favor de la posición de Harsanyi. Después propongo una modificación a una de las premisas de Rawls y, apoyado en mi argumentación sobre los atributos positivos de maximin, concluyo a favor de Rawls desde esta visión ampliada. Finalmente, reviso la situación actual del debate y una interpretación reciente sobre la interacción entre las premisas morales y las propuestas de justicia de Rawls y Harsanyi.

### **31. LO MORAL Y LO RACIONAL EN RAWLS**

Argumento que la esencia para poder concluir sobre esta disputa está en una adecuada comprensión del papel que juegan lo racional y lo moral en Rawls. Es importante tener presente que para Rawls el artificio de la posición original es parte de la búsqueda de un fundamento moral razonable que tenga raíces en nuestra concepción de nosotros mismos, y en nuestra relación con la sociedad. Así, el objetivo de esta indagación moral es encontrar una fundamentación razonable para aceptar y acordar una concepción de justicia adecuada. La situación inicial de Rawls le permite modelar dos ideales morales concretos: por un lado, la noción de imparcialidad, que es activada a través del artificio del velo de ignorancia, y por otro, la igualdad, que se funda en la autonomía de cada individuo, que tiene la facultad de participar al mismo nivel del resto de los ciudadanos, a través de las partes representantes, en el proceso de toma de decisión sobre los principios de justicia. Sin embargo, a pesar de las fuertes condiciones morales de la posición original, Rawls reconoce que los seres humanos tenemos ideas e intuiciones acerca de la justicia que son conflictivas entre sí, y por tanto los argumentos de tipo moral que surgen de esas intuiciones particulares serán siempre poco convincentes para algunas personas.

Por tanto, en la búsqueda de un acuerdo razonable, Rawls se ve obligado a apoyarse en un criterio de justificación con mayor validez universal que lo moral: la

elección racional. Así, al utilizar la posición original y sus condicionantes, como el velo de ignorancia, el autor supone solamente que las personas representantes ahí situadas son capaces de hacer una elección racional: actúan motivados para promover sus intereses personales. Si la justificación de los principios es derivada de la noción de racionalidad, estos principios pasarán la prueba de afinidad, que implicaría que personas reales, con sus propios intereses personales, podrán identificarse con la racionalidad de los principios. Podemos decir que, en una sociedad diversa, la fuerza justificatoria de la elección racional es el punto de apoyo común.

Sin embargo, no basta con que la elección sea racional, y no todas las elecciones racionales podrían constituir una concepción de justicia. Además de la prueba de afinidad, los principios de justicia elegidos en la posición original deben ser reconocidos como principios de justicia. Para ser principios de justicia deben pasar la prueba de verificación, en el sentido de que cada persona de la sociedad real puede confirmar que los principios han sido elegidos desde un punto de vista imparcial moral, que refleje lo que la sociedad considera correcto y aceptable para sus miembros, y que serían elegidas por un legislador imparcial.

Así, la posición original tiene dos conexiones, una hacia el punto de vista moral del otro, y la otra hacia los individuos racionales reales. Las dos conexiones son esenciales, y es la combinación de las dos lo que da a la posición original su fuerza.

## **32. EN LA BÚSQUEDA DE UNA RACIONALIDAD MÁXIMA**

En la búsqueda de un nivel máximo de racionalidad, Rawls postula ambiciosamente que la justicia es parte, “quizá la más significativa”, de la teoría de decisión racional. Hasta ese momento nadie se había atrevido a postular el rasgo racional de la justicia de esa manera tan contundente. La pregunta que surge es ¿por qué Rawls pretende dar a la justicia ese grado tan elevado de racionalidad? El filósofo americano ha definido desde hace mucho tiempo antes de la publicación de *TJ* que su objetivo es desbancar al utilitarismo, y ha estudiado a fondo esta teoría y a sus mejores exponentes, entre ellos a Harsanyi.

Me parece que Rawls sabe con certeza que su opositor, el utilitarismo, es una teoría ética que se ha desarrollado hasta convertirse en una teoría económica sofisticada,

cuyos fundamentos han sido solidificados con herramientas propuestas por la teoría de decisión racional. Esto es evidente cuando Rawls escribe:

durante mucho tiempo la teoría sistemática predominante en la filosofía moral moderna ha sido alguna forma de utilitarismo. Una razón de ello es que ha sido defendida por una larga serie de escritores brillantes que han construido una doctrina intelectual verdaderamente impresionante en sus alcances y en su refinamiento...los grandes utilitarios...eran teóricos sociales y economistas de primera línea, y la doctrina moral que elaboraron pretendía satisfacer las necesidades de sus más vastos intereses y ajustarse a un esquema general.<sup>104</sup>

Así, argumento que cuando Rawls emprende su objetivo, desea estar a la par del utilitarismo en el nivel de racionalidad de su teoría. Para ello tiene dos alternativas. La primera es argumentar y fundamentar la racionalidad de las partes representantes en la situación original desde todos los ángulos posibles, lo cual lleva a cabo a lo largo de todo el texto, exponiendo argumentos que incluyen el término 'racional' en más de quinientas ocasiones. Sin embargo, ante la duda de que hubiera alguna fisura que pudiera hacer cuestionable la racionalidad de su sistema, pienso que Rawls va un paso más allá y busca cerrar la posibilidad de cualquier grieta postulando que 'la justicia es una parte, quizá la más significativa, de la teoría de la elección racional'.

Pero ¿qué es exactamente la teoría de la elección racional y qué ve Rawls en ella que le atrae, más allá del beneficio de aumentar el peso de la racionalidad de su propuesta? A grandes rasgos, se puede decir que la teoría de decisión, como marco teórico que se utiliza para comprender y modelar de una manera formal el comportamiento social y económico, tiene cuatro principios fundamentales<sup>105</sup>: primero, los individuos son racionales y piensan de una manera lógica; segundo, los individuos tienen intereses; diferentes personas tienen diferentes intereses, y los intereses de la persona definen su utilidad; tercero, los individuos hacen elecciones, y estas elecciones afectan su utilidad recibida; cuarto, los individuos hacen elecciones racionales con el objetivo primordial de aumentar sus utilidades, aun cuando las personas no tienen

---

<sup>104</sup> *Op. Cit.*, p.9

<sup>105</sup> [http://www.colorado.edu/Sociology/Mayer/Contemporary%20Theory/Rational%20Choice\\_files/frame.htm](http://www.colorado.edu/Sociology/Mayer/Contemporary%20Theory/Rational%20Choice_files/frame.htm)

control completo sobre los resultados de sus elecciones. A partir de estas nociones sencillas, se ha construido un sofisticado edificio conceptual que ha sido usado exitosamente en dos áreas: en la economía, el terreno de Harsanyi, y posteriormente en las ciencias sociales. Para una definición más formal, acudiré a Michael Resnik, quien dice:

La teoría de decisión es el producto de los esfuerzos conjuntos de economistas, matemáticos, científicos sociales, y especialistas en estadística para la mejor comprensión acerca de cómo los individuos y los grupos toman o deberían tomar decisiones.<sup>106</sup>

Regresando por un momento a la definición de la posición original, encuentro que hay algunos elementos en la teoría de decisión que resultan familiares. Por ejemplo, las partes representantes son racionales, “y están interesadas en promover sus propios intereses”.<sup>107</sup> Asimismo, cuando se especifica la postura de las partes respecto de los beneficios, las cuales “preferirán tener más bienes sociales primarios a tener menos”.<sup>108</sup> Así, encuentro evidencia de que Rawls no solamente apela a la pertenencia de la justicia a la teoría racional en busca de una racionalidad superior, sino que además ha tomado los conceptos básicos de la teoría de elección racional para el diseño de su posición original.

Esta sospecha se reafirma con lo que Resnik escribe al respecto: “Rawls y Harsanyi han tomado prestado de la teoría de decisión para argumentar a favor de ciertas posiciones en filosofía social.”<sup>109</sup>

Concluyo así, que para Rawls la mejor manera de completar el trabajo de posicionar su teoría con un grado máximo de racionalidad es a través de una estructura formada por tres ejes. El primero es la argumentación abundante sobre la presencia de lo racional en todo el texto. El segundo es tomar prestados elementos conceptuales

---

<sup>106</sup> Michael Resnik, *Choices: An Introduction to Decision Theory*. EUA: University of Minnesota Press, 1987, p.3

<sup>107</sup> *Op. Cit.*, p.24

<sup>108</sup> *Op. Cit.*, p.140

<sup>109</sup> *Op. Cit.*, p.43.

fundamentales de la teoría de elección racional para diseñar la racionalidad de los agentes en la posición original. El tercero y último, es este esfuerzo ambicioso de evitar cualquier fractura de racionalidad en su propuesta que lo lleva a postular que la justicia, y por tanto su sistema, es parte 'quizá la más significativa', de la teoría de elección racional.

### **33. EL ERROR DE RAWLS**

Pienso que Rawls comete un error al hacer esa pretensión de máxima racionalidad. Por un lado, me parece que la teoría de decisión es una metodología lógica y técnica, cuyo avance se ha fundamentado en la formulación de algunos principios y axiomas que le han ido permitiendo atacar problemas de elección racional en esquemas cada vez más complejos. La justicia, por su parte, es un tema de un alto contenido moral que no puede pertenecer a una metodología técnica. La pretensión de Rawls en este sentido resulta a primera vista como algo incongruente, pues lo moral nunca podrá ser parte 'significativa' de una disciplina técnica.

Por otro lado, al presentar a la justicia como parte de la teoría de la elección racional, Rawls invita la reacción crítica de los expertos en el tema. Un caso conocido es el de Brian Barry, que ya fue comentado. Otro es el caso de Harsanyi, quien no sólo conoce el tema, sino que ha sido uno de los contribuidores a una de sus ramas: la teoría de juegos. Además, Harsanyi se ha apoyado en las fuertes ligas que hay entre la economía y la teoría de elección para algunos de sus desarrollos sobre el utilitarismo.

Por último, con esta pretensión Rawls se obliga a sí mismo a defender su principio maximin desde el terreno de la teoría racional. Al colocarse en el campo de esta disciplina, él se encuentra sin argumentos de peso técnico pues no es un experto en la materia -mientras que Harsanyi sí lo es- y se ve obligado a recurrir a defensas escurridizas como las esgrimidas en su respuesta a la publicación crítica de Harsanyi. En la sección 29 expuse los elementos de la réplica de Rawls. Su primer argumento es que "el criterio maximin no está hecho para ser aplicado a situaciones de pequeña escala, como es el caso de cómo un doctor debiera tratar a sus pacientes... maximin es un criterio macro, no micro"<sup>110</sup>. Otorgo la razón a Harsanyi cuando contesta que los

---

<sup>110</sup> *Op. cit.*, p.142

contraejemplos que ha presentado en su reprobación no tienen relación con la escala y cuando pregunta, muy agudamente, si para Rawls existe un valor  $x$ , a partir del cual las circunstancias de la justicia dejan de ser de pequeña escala, y por tanto de tamaño suficiente para merecer la aplicación del principio maximin. Pienso que Harsanyi acaba de un plumazo con este argumento defensivo de Rawls, quien no vuelve a esgrimir ese razonamiento. Otro argumento de Rawls es que maximin requiere mucha menos información para ser aplicado que la maximización de la utilidad, que requeriría de las curvas de utilidad de cada miembro de la comunidad. Este razonamiento me parece también debatible, dado que Rawls no provee en su sistema un método práctico para determinar cuáles miembros de la sociedad constituyen el sector de los 'menos favorecidos'; no hay una formulación concreta de la manera de calcular un índice de bienes primarios. El tercer argumento refiere a la necesidad de que el criterio de justicia sea público y relativamente fácil de comprender por las personas de todos los sectores sociales. Ciertamente parece más sencillo lograr que el gran público entienda los dos principios rawlsianos que el utilitarismo, pero pienso que esto no es un elemento que sostenga la propuesta de Rawls de concluir que las partes en la posición original llegarán de una manera racional a la elección de maximin por encima del utilitarismo. Otra justificación a favor de maximin es el que los principios deben ser tolerables de manera permanente para todos los miembros de la comunidad. Ciertamente maximin tendrá más atractivo intuitivo para los sectores más amplios de la sociedad que el utilitarismo, pero, nuevamente, esto no me parece un respaldo a la racionalidad de la decisión en la posición original. Finalmente, Rawls acude a la mayor capacidad de maximin para permitir a los ciudadanos alcanzar una personalidad libre e igual. Pienso que el criterio maximin es más igualitario que el de la utilidad, pero ello no suma en el vector de racionalidad de la decisión. En sinopsis, los atributos de maximin destacados por Rawls en su defensa de maximin son importantes pues tienen un peso moral significativo para el bienestar de la sociedad. Sin embargo, esos atributos morales no sirven para respaldar la afirmación de que maximin será elegido por los agentes estrictamente racionales por encima de la utilidad en la situación inicial.

Así, en la búsqueda de un sostén racional máximo, Rawls teje las redes de su propia trampa al proponer que su sistema y la justicia sean parte de la elección racional. Ello tiene tres consecuencias negativas: primero, la incongruencia lógica de pretender insertar un tema moral dentro de una disciplina técnica; segundo, invitar la reacción crítica de los expertos en teoría normativa de elección; y tercero, restringir el debate al terreno técnico de la decisión racional, en el que no aplican los atributos morales de maximin. Como resultado de todo lo anterior, la postura de Harsanyi es avalada por la mayoría de los concedores de la elección racional, quienes dan sus argumentaciones desde la óptica de esta disciplina.

#### **34. ARGUMENTACION A FAVOR DE MAXIMIN**

La primera tarea que debe emprender Rawls para vencer al utilitarismo y declarar ganadores a sus dos principios es doble: por un lado, debe proveer argumentos para prohibir la posibilidad de que apliquen las probabilidades, y en particular la equiprobabilidad, a la situación inicial. Por el otro, debe proponer condiciones para argumentar que, cuando no aplican las probabilidades, la elección de principios de justicia desde la situación inicial deberá ser guiada por la regla maximin, en el proceso de elegir la mejor concepción de justicia para la sociedad a surgir. Empiezo por el primer objetivo.

El enemigo de Rawls es el utilitarismo. Harsanyi ha demostrado que, a partir de la posición original rawlsiana, personas racionales elegirán el principio de la utilidad como principio de justicia, con la condición de que aplique el principio de razón insuficiente de Laplace, en virtud del cual en una situación de incertidumbre total, en la que no hay información para estimar probabilidades concretas, las personas que harán la elección necesariamente harán uso de las probabilidades subjetivas; en ese caso se asigna una probabilidad igual a cada posible alternativa. Esta condición es conocida como equiprobabilidad, y es uno de los elementos fundamentales de la teoría Bayesiana. Dado que la premisa de equiprobabilidad lleva a las partes racionales a la elección de la utilidad, y no a la de maximin, el primer opositor de fondo de Rawls es la premisa de equiprobabilidad.

Rawls entiende esto, y diseña la posición original con unas especificaciones tales, que no haya lugar para la equiprobabilidad, ni cualquier otro tipo de asignación de probabilidades. El dispositivo del velo de ignorancia cumple esta función, pues la falta de toda información acerca de la naturaleza probable de la sociedad, y del lugar que ocuparán en ella, no permite a los agentes racionales asignar ningún tipo de probabilidad, y menos una probabilidad igual a cada posible arreglo institucional y a cada posible lugar dentro de la sociedad.

Escribe Rawls:

puede resultar sorprendente que el significado de probabilidad surja como un problema de filosofía moral, en el campo de la teoría de la justicia. Sin, embargo es una consecuencia inevitable de la doctrina contractual, la cual concibe la filosofía moral como parte de la teoría de la elección racional.<sup>111</sup>

Es evidente, pues, que las probabilidades surgen en el campo de la justicia por la decisión de Rawls de tomar elementos de la teoría normativa de la decisión en su diseño de la posición original, la cual considera las probabilidades. Él sabe que tiene que lidiar con este concepto matemático del azar, por ser parte del contexto de elección racional en el que ha elegido insertar su sistema.

Harsanyi argumenta en favor del uso de la premisa de equiprobabilidad en la posición original apoyado en el hecho de que “la teoría Bayesiana de decisión muestra con argumentos matemáticos rigurosos que cualquier tomador de decisiones cuyo comportamiento es consistente con unos pocos postulados de racionalidad simplemente no puede evitar actuar como si usara las probabilidades subjetivas.”<sup>112</sup> Esto significa que en el fondo de la postura de Harsanyi sobre la necesidad de la equiprobabilidad hay un argumento técnico matemático, y no un razonamiento moral. Sin embargo, Harsanyi en el mismo texto<sup>113</sup> da un respaldo moral adicional a su supuesto clave cuando dice: “la premisa de equiprobabilidad se puede justificar...también por el principio moral de asignar el mismo peso *a priori* a los

---

<sup>111</sup> *Op. cit.*, p.165

<sup>112</sup> *Op. cit.*, p.47

<sup>113</sup> *Idem.*, p.48

intereses de cada individuo.” Interpreto que esto significa que para Harsanyi la situación inicial es tal que hay una impersonalidad que no permite distinguir a unas personas de otras. Volveré a este rasgo más adelante.

Coincido con Rawls cuando argumenta que no es posible aplicar ninguna estimación de probabilidad, dado el elevado grosor del velo de ignorancia. Argumento que intuitivamente no es plausible asignar una igual probabilidad a cada alternativa de arreglo social y a cada posición dentro de la comunidad, cuando la incertidumbre es total, aun cuando exista un argumento matemático que así lo proponga. Así, pienso que Rawls logra superar el primer obstáculo y convencer al lector de que no aplican las probabilidades.

Avanzo ahora al segundo reto de Rawls: si aceptamos que las probabilidades no aplican a la posición original, ¿podemos concluir que las partes racionales elegirán necesariamente el principio de justicia maximin? Este punto ha sido un tema de intensa discusión entre los teóricos de la decisión. El consenso general es que maximin sí es una posible elección racional si no aplican las probabilidades, aunque, de hecho, como Binmore<sup>114</sup> y otros han señalado, no es la única solución. Para efectos de este trabajo, consideraré que el criterio maximin es una solución racional aceptable.

En resumen, me parece que Rawls hace un caso plausible de la no aplicación de las probabilidades a la posición original, y de la validez de maximin como una posible solución a la elección de principios de justicia en la posición original, aunque no es la única. Procedo ahora a revisar los otros argumentos esgrimidos por Rawls para la defensa de su criterio. Me parece relevante que al autor acepte abiertamente algunos inconvenientes de maximin cuando dice que este criterio no es, en general, una guía apropiada para elegir bajo condiciones de incertidumbre. Sin embargo, como ya expliqué en la sección 18, la regla es atractiva en determinadas situaciones caracterizadas por ciertos rasgos especiales.

Por tanto, pareciera que el mero rechazo de las probabilidades no es, a los ojos de Rawls, una condición suficiente para adoptar la estrategia maximin, y que es necesario presentar más razonamientos a favor de este criterio. En efecto, Rawls argumenta que

---

<sup>114</sup> *Op. cit.*, p. 107

además de esta primera condición de la no aplicación de probabilidades se requiere una segunda condición: “la persona que escoge tiene una concepción del bien tal que le importa muy poco o nada lo que pueda ganar todavía por encima del mínimo estipendio que seguramente obtiene al seguir la regla maximin”.<sup>115</sup> En esencia esto significa que no conviene correr el riesgo de obtener más ventajas, sobre todo si existe la posibilidad de quedar colocado por debajo del mínimo aceptable. Una tercera y última condición es que “las alternativas rechazadas tienen resultados que difícilmente pueden aceptarse”<sup>116</sup>, aludiendo al grave riesgo de elegir otros sistemas que pudieran, por ejemplo, sacrificar la libertad en aras de mayores beneficios económicos y sociales, como podría ser el caso del utilitarismo.

Rawls sostiene que la presencia de estos tres rasgos en la situación inicial justifica la selección de una estrategia maximin por las partes cuando éstas eligen entre los principios de justicia y el principio de la utilidad.

### **35. OBJECIONES A LA CRÍTICA DE HARSANYI**

La postura de Harsanyi ha gozado del apoyo de la mayoría de los estudiosos de la disputa, quienes han sido primordialmente especialistas de la teoría de elección racional. Hasta cierto punto, no es sorprendente el veredicto favorable de sus colegas. Sin embargo, estos comentaristas de la decisión racional con frecuencia han pasado por alto las posibles objeciones al razonamiento de Harsanyi.

Mi inconformidad con la crítica de Harsanyi a maximin, y en particular con el contraejemplo de la elección entre el trabajo en Nueva York con bajo salario -sin necesidad de cambiar de ciudad-, y el trabajo mejor pagado en Chicago -que requiere tomar un vuelo que podría estrellarse-, es que éste no responde al argumento de Rawls en el sentido de que maximin es una regla apropiada de decisión solamente bajo las condiciones cuidadosamente definidas de extrema incertidumbre, en las que no pueden aplicar las probabilidades. Me parece que Harsanyi únicamente demuestra que maximin es una regla inapropiada para elecciones bajo circunstancias en las cuales las

---

<sup>115</sup> *Op. cit.*, p.151

<sup>116</sup> *Idem*

personas tienen mucha información acerca de las probabilidades, aunque la probabilidad precisa de cada resultado no es conocida. Por tanto, pienso que la objeción de Harsanyi no aborda, y mucho menos refuta, los argumentos de Rawls referidos a que maximin es el criterio apropiado de elección racional bajo condiciones de incertidumbre extrema.

Argumento que la objeción de Harsanyi está motivada por la visión de que el criterio de elección racional no debería cambiar en ninguna situación, con la pretensión de establecer que el único criterio válido para situaciones de incertidumbre es la maximización de la utilidad, a partir de la premisa de equiprobabilidad.

### **36. CONCLUSIÓN DESDE LA OPTICA RACIONAL ESTRICTA**

Pienso que si Rawls hubiera mantenido las premisas de racionalidad de los agentes en el diseño de la Posición Original, sin llegar al extremo de postular que la justicia es parte de la teoría de elección racional, le podría haber sido posible hacer una defensa de su principio maximin frente al principio de utilidad, fuera del terreno de la aplicación de la teoría normativa de decisión. Desafortunadamente Rawls insiste en varias ocasiones en esta relación fundamental, lo que le lleva a que la mayoría de los comentaristas de la metodología de la elección racional esté a favor de Harsanyi.

Encuentro que Rawls, en su búsqueda de máxima racionalidad, postula dos premisas contradictorias, generando un conflicto lógico que no tiene solución. Al insertar su sistema en la teoría de decisión racional, la propia normatividad de esa disciplina hace mandatoria la aplicación de la equiprobabilidad, como lo lleva a cabo Harsanyi. Paralelamente, Rawls busca evitar a toda costa la aplicación de la equiprobabilidad, pues sabe que esta premisa conduce inevitablemente al principio de justicia de la utilidad; para ello diseña un velo de ignorancia mucho más grueso que el de Harsanyi, de manera que en la posición original no sea posible hacer estimación alguna de probabilidades de alternativas de sistemas sociales ni sobre los diferentes lugares que los representantes podrían ocupar en la escalera social.

Por tanto, concluyo que el principio maximin, tal como la planteó su autor, no es superior al principio de la utilidad, si nos situamos estrictamente en la perspectiva de la aplicación correcta de la teoría de la decisión racional. Esto implica que la regla

maximin no es un principio que sería elegido por agentes racionales en la posición original, por encima del principio de la utilidad, para sopesar las diferentes concepciones de justicia y elegir la más satisfactoria. Dado que maximin no sería seleccionado por las partes, y el utilitarismo sería preferido, concluyo que la teoría de justicia de Rawls no se sostiene frente a la crítica de Harsanyi. Así, queda contestada la pregunta principal de este trabajo de investigación.

### **37. ¿CÓMO SALVAR LA TEORÍA DE JUSTICIA DE RAWLS?**

Me parece que la propuesta rawlsiana es muy poderosa y que el argumento maximin cumple funciones morales muy amplias. Adicionalmente, pienso que la perspectiva de los comentaristas que han descartado maximin -y por ende toda la propuesta rawlsiana- ha sido ciertamente estrecha, limitándose a la aplicación estricta de la teoría normativa de la elección. Esto me lleva a la pregunta ¿existe alguna manera de salvar la TJ de Rawls? Pienso que la validez de la teoría de Rawls puede ser rescatada si retiramos una sola premisa, la más ambiciosa y la menos justificada, que refiere a la pretensión de que la justicia es posiblemente la parte 'más significativa' de la teoría de elección racional.

Así, propongo una versión modificada del sistema de Rawls, en la cual se mantengan las condiciones de diseño de la posición original, lugar desde el cual se elegirán los principios de justicia. Los agentes tomadores de decisión tendrán el mismo grado de racionalidad, el velo de ignorancia mantendrá su mismo grosor. El objeto de la justicia seguirá siendo la estructura básica social, las alternativas a considerar para la elección del procedimiento de justicia serán las mismas, incluyendo desde luego los principios de Rawls y el utilitarismo. El resto de las condiciones en la situación inicial se conservan. Solamente se elimina la fuerte restricción de la pertenencia a la teoría normativa de elección

Argumento que a partir de esta modificación, aparentemente sutil, pero de gran calado, ya no es estrictamente necesario situar el debate Rawls-Harsanyi en el terreno de la teoría de la decisión racional, y por tanto, ya no será indispensable el uso de las premisas propias de esta disciplina, como la equiprobabilidad. Con esta alteración, que no modifica la fuerza del planteamiento rawlsiano, ya es posible salir del conflicto de las

dos premisas contradictorias. Ya es plausible proponer un velo de ignorancia tan voluminoso, que genera una ignorancia total sobre los rasgos de los electores y sobre el futuro, de manera que no es posible considerar la aplicación de ninguna probabilidad, sin entrar en conflicto con la metodología de la elección racional.

Con esta alteración liberadora que elimina la contradicción lógica de dos premisas en conflicto, haré una nueva evaluación comparativa del criterio maximin como lo concibe Rawls y del utilitarismo como lo presenta Harsanyi, en base a una revisión de los atributos y debilidades de cada uno.

Empiezo con el criterio de la utilidad de Harsanyi, que postula dos elementos: la crítica a maximin, y la defensa de la utilidad. Respecto a la reprobación de maximin, ya expuse mis objeciones en el sentido de que sus contraejemplos no dan una verdadera refutación a la postura de Rawls, pues no invalidan la elección de maximin bajo las tres condiciones estipuladas. En relación a las virtudes del principio utilitario, el argumento principal de Harsanyi es la necesidad ‘matemática’ de la aplicación de la equiprobabilidad a la posición original que, como ya expuse, queda en duda dada la máxima incertidumbre de la situación inicial, en la que no es posible asignar ningún tipo de probabilidad.

Reviso ahora los atributos de maximin. Me parece que las críticas más conocidas a este principio, por parte de los comentaristas de la elección racional hace ya más de treinta años, han logrado dejar asentado en la comunidad académica que la regla simplemente no tiene sustento racional, dado que el argumento de la equiprobabilidad es favorecido por esos críticos. Soy de la opinión que esto ha provocado que se pasen por alto los atributos más importantes de maximin, que expongo a continuación.

Propongo que maximin es más robusto de lo que se ha reconocido, y que este argumento lleva a cabo funciones importantes en la clarificación de la naturaleza y las implicaciones del razonamiento en el sistema de justicia rawlsiano. Argumento que el mínimo satisfactorio que propone Rawls no es una cantidad mínima de una canasta de bienes primarios, sino una “concepción mínima adecuada de la justicia”<sup>117</sup>; es decir, una concepción que provee la garantía mínima más satisfactoria de la protección de los

---

<sup>117</sup> *Op. Cit.*, p.169

derechos fundamentales de las personas. Los encargados de elegir seleccionarán una teoría que minimice la invasión de sus derechos fundamentales, promueva una oportunidad igual para desarrollar y explotar sus talentos naturales, y mitigue las desigualdades que existan en un orden social que asegure la igualdad de oportunidades. Este mínimo satisfactorio es asegurado por los dos principios en su orden lexicográfico. Pienso que los individuos en la posición original no se concentran en la forma en que serán asignados los bienes primarios, sino en valorar el carácter de las distintas concepciones de justicia, en los diferentes intereses que cada una protege, y en el tipo de mundo social y político que surgiría a partir de ellas. Al elegir el principio maximin, hacen uso de esta regla para una tarea muy concreta: la evaluación de la substancia y las implicaciones prácticas de la amplia variedad de concepciones propuestas.

Así, los electores utilizan el criterio maximin para seleccionar los dos principios rawlsianos porque constituyen la concepción mínima de justicia más satisfactoria dado que permiten la más satisfactoria protección mínima de los intereses de la libertad, y porque les aseguran su prioridad por encima de cualesquiera otros intereses fundamentales. Es la concepción de los dos principios, tomados como un todo, lo que constituye el mínimo satisfactorio; los requerimientos de los principios elegidos están atados unos con otros como una concepción de justicia que aplica a la estructura básica social como un todo, y el hecho de que esta concepción asegura un mundo político y social satisfactorio -y no meramente un lote de bienes, recursos y protecciones- es crucial para el razonamiento que lleva a maximin. Me parece que esta función trascendental de maximin no ha sido identificada ni valorada por los teóricos de la elección racional, quienes se han concentrado en la obligada aplicación de las probabilidades.

### **38. CONCLUSIÓN DESDE UNA ÓPTICA AMPLIADA**

Con el retiro de la restricción metodológica de la equiprobabilidad, y la consideración de esta visión ampliada de la fuerza de maximin, concluyo que el principio maximin es la elección que harán las partes situadas en la posición original, cuyos rasgos quedan especificados por las tres condiciones propuestas por Rawls. Maximin será elegido por

encima del criterio de la utilidad, el cual no puede brindar una concepción mínima adecuada de justicia en la que se asegure la protección mínima de los derechos más fundamentales de las personas.

El peso moral de maximin queda establecido por su capacidad para constituir una idea integral mínima de justicia que cumpla con la condición estricta de Rawls que exige dar una prioridad definitiva a la libertad, por encima de cualesquiera otros bienes primarios.

### **39. LA DISPUTA RAWLS-HARSANYI EN LA ACTUALIDAD**

Han pasado más de cuarenta años desde la publicación de la crítica de Harsanyi. En este período tanto Rawls como su contrincante fueron modificando algunas de sus premisas, y algunos comentaristas han ofrecido nuevas y más amplias perspectivas del debate, por lo que resulta relevante comentar al respecto.

Ante la abundante crítica acerca de la fallida racionalidad de la derivación de maximin, Rawls hizo dos ajustes correctivos a su teoría. Por un lado, propuso en su obra *Liberalismo Político*, publicada en 1993, en una discreta nota, un cambio fundamental en la relación entre la justicia y la teoría de decisión:

Aquí corrijo un comentario en *Teoría...* donde dice que la teoría de justicia es una parte de la teoría de la decisión racional....esto es simplemente incorrecto. Lo que se debería haber dicho es que el proceso de dar cuenta de las partes, y de su razonamiento, usa la teoría de la decisión racional, pero solamente en una forma intuitiva. Esta teoría es ella misma parte de una concepción política de la justicia, que intenta dar cuenta de principios razonables de justicia. No existe una intención de derivar esos principios del concepto de racionalidad como el único concepto normativo. Creo que el texto de *Teoría* como un todo apoya esta interpretación.<sup>118</sup>

Así, Rawls finalmente niega la relación íntima entre la justicia y la teoría normativa, que había propuesto en su *TJ* original. Creo que este cambio del autor es propuesto más como una salida para dar fin a las críticas de los especialistas de la decisión, que como una reformulación de fondo de su teoría, como yo lo postulo.

---

<sup>118</sup> John Rawls, *Political Liberalism*, expanded edition. New York: Columbia University Press, 2005. Nota p.53

Por otro lado, en su obra *Justicia como Equidad, Una Reformulación*<sup>119</sup>, publicado en 2001, un año antes de su fallecimiento, Rawls da un giro significativo adicional cuando indica que

no es esencial que las partes usen la regla maximin en la posición original. No es más que una herramienta heurística. Centrarse en los peores resultados tiene la ventaja de forzarnos a considerar cuáles son realmente nuestros intereses fundamentales cuando se trata de diseñar la estructura básica.<sup>120</sup>

Esta modificación implica que maximin ya no es indispensable para la justificación del principio de diferencia en la posición original. En esta última versión de su sistema, Rawls apela a la justificación moral del argumento maximin cuando refiere al papel de las partes como representantes de otros individuos iguales, y su responsabilidad como tales. Dice Rawls

Sólo de este modo (seleccionando los dos principios de justicia) pueden (los representantes) obrar responsablemente como fideicomisarios: esto es, proteger de forma efectiva los intereses fundamentales de la persona a la que cada cual representa y, al mismo tiempo, asegurarse de que evitan posibilidades cuya realización sería absolutamente intolerable.<sup>121</sup>

Así, la elección del uso del principio maximin no obedece estrictamente a una deducción racional. Su justificación es moral, pues permite a las partes cumplir con su compromiso como representantes de otros a los que debe proteger en su elección de los principios de justicia.

Harsanyi, a su vez, hace una muy importante modificación ocho años más tarde. De hecho, el Nobel acepta su equivocación respecto de la premisa más relevante en el argumento esgrimido para refutar la propuesta rawlsiana de maximin, y abandona el planteamiento sobre la aplicación de la equiprobabilidad, como una necesidad

---

<sup>119</sup> John Rawls, *Justicia como Equidad, Una Reformulación*. Barcelona: Paidós, 2002.

<sup>120</sup> *Ibidem*, p.141

<sup>121</sup> *Ibidem*, p.144.

‘matemática’ imprescindible en situaciones de incertidumbre. En efecto, escribe el economista

En su forma tradicional, el principio de indiferencia, también llamado principio de razón insuficiente, afirma que cuando las condiciones permiten dos o más resultados diferentes, mas no hay evidencia para favorecer un resultado particular sobre cualquier otro, debemos asignar la misma probabilidad a cada resultado. En esta forma, el principio es demasiado vago para tener un uso real. También está sujeto a la importante objeción lógica de que intenta alcanzar una conclusión positiva (la de equiprobabilidad) a partir de la mera ignorancia (debido a la ausencia de información que favorezca a algún resultado en particular) -un intento que no puede tener éxito.<sup>122</sup>

Este es un giro sumamente radical en la postura de Harsanyi y de los teóricos de la decisión, pues anteriormente la equiprobabilidad era un dogma intocable. Ahora esta noción puede verse como una premisa que puede o no ser aplicada a la situación inicial.

Recientemente el profesor estadounidense Moehler<sup>123</sup> publicó un ensayo que toma en consideración los cambios de premisas en Rawls y Harsanyi, para dar una nueva interpretación al debate. En esencia, Moehler argumenta que la teoría de justicia de cada uno de los pensadores modela ideales morales distintos. Bajo esta óptica, a los ojos de Moehler, la propuesta de Harsanyi plasma tres ideales: imparcialidad, igualdad e impersonalidad. La imparcialidad es instalada por el dispositivo del velo de ignorancia, un velo más delgado que el de Rawls pues los individuos en la situación original conocen las circunstancias específicas de la sociedad, y los perfiles de preferencias de todos los miembros; mas no conocen los nombres de las personas que están ligadas a cada perfil de preferencias. La igualdad es impuesta a través de la premisa de equiprobabilidad, la cual establece que cada persona tendrá la misma posibilidad de ocupar cualquiera de las posiciones sociales. El modelo de Harsanyi integra también la idea de impersonalidad, por la que los individuos no podrán decidir

---

<sup>122</sup> John Harsanyi, “Bayesian Decision Theory, Subjective and Objective Probabilities, and Acceptance of Empirical Hypotheses”, en *Synthese* 57, 1983, p. 352.

<sup>123</sup> Michael Moehler, “The Rawls–Harsanyi Dispute: A Moral Point of View” en *Pacific Philosophical Quarterly*, doi: 10.1111/papq.12140. 2015.

sobre cuestiones de justicia basados en sus propios intereses, o en los intereses personales de los otros, sino solamente basados en un interés público socialmente construido. La impersonalidad es el sustento moral nuclear del utilitarismo y es, para Rawls, la principal razón por la que el utilitarismo no toma en cuenta la individualidad y la autonomía de las personas.

El modelo de Rawls, en el análisis de Moehler, propone los ideales morales de imparcialidad, igualdad, autonomía, reciprocidad y estabilidad de cooperación. La imparcialidad es instalada a través del velo de ignorancia, más grueso que el de Harsanyi; desde esta situación, las partes no podrán favorecer ninguna posición particular injustificadamente. La posición original también modela la igualdad, puesto que todos los miembros de la sociedad tienen el mismo peso en el consenso total que se deberá lograr para la aprobación de los principios. La autonomía, distintivo moral primordial de la propuesta de Rawls, está incluida en el diseño de la posición original, que garantiza a los individuos la protección de su autonomía como personas separadas. Adicionalmente, el principio de diferencia, debido a la consideración del rigor de lo que significa comprometerse<sup>124</sup>, modela adecuadamente el ideal de una sociedad como un sistema de cooperación equitativo entre ciudadanos libres e iguales, los cuales comparten el ideal moral de la reciprocidad y la preocupación por la estabilidad de la cooperación.

Entendidas desde esta óptica, cada teoría de justicia postula ideales morales particulares. Así, las partes representantes situadas en la posición original definirán primeramente cuáles son los ideales morales con los que comulgan, de entre los que promueven los dos sistemas. Si las partes se identifican mayormente con la autonomía y los otros ideales de la justicia rawlsiana, elegirán el criterio maximin y llegarán a los dos principios de justicia. Por el contrario, si los representantes están más vinculados con los valores morales del utilitarismo, incluida la impersonalidad, optarán por la regla de la utilidad y seleccionarán el utilitarismo. En este sentido, entendida la situación inicial como dos modelos con premisas morales diferenciadas, podemos decir, argumenta Moehler, que no hay vencedor ni vencido. Si somos kantianos y damos un

---

<sup>124</sup> Traducción de la expresión en inglés 'strains of commitment'.

peso relevante a la autonomía de la persona como un valor moral superior, elegiremos los principios de Rawls. Encuentro difícil de justificar el no darle peso a la autonomía del individuo, pero si se parte de la premisa moral de la impersonalidad, el principio elegido desde la situación original será el utilitarismo.

Me parece que el valor de esta nueva mirada hacia el debate es el énfasis en los ideales morales que cada concepción de justicia busca presentar, y el dejar claro que la elección de los principios de justicia no puede ser justificada exclusivamente por métodos analíticos de decisión. En todo caso, esta propuesta manifiesta la vigencia del debate, y ayuda a profundizar y enriquecer nuestra comprensión de los requerimientos de la justicia.

#### **40. COMENTARIOS FINALES**

John Rawls fue uno de los filósofos morales y políticos más importantes del siglo XX, y será obligadamente tomado en cuenta en cualquier discusión filosófica futura sobre la justicia.

Su teoría brindó nuevas maneras de entender la posibilidad de una 'sociedad bien-ordenada' que fuese perfectamente justa, inspirada ciertamente en las referencias del sistema social norteamericano, y podría decirse que su modo de reflexión está contextualmente demarcado en términos de la historia de esa cultura política.

Rawls dedicó su pensamiento no al servicio de lo verdadero sino de lo práctico, y su filosofía no concierne a los problemas universales sino a los del ser humano en su ámbito sociopolítico concreto. Rawls resucitó el modelo contractualista para hacer filosofía política, y para ello se sirvió de las herramientas analíticas desarrolladas por la teoría de la decisión racional que en ese momento estaban ya acreditadas. Asimismo, recuperó para la justicia un papel preponderante dentro de la filosofía política que no había conocido desde la época de Platón y Aristóteles. También brindó para la vida política un paradigma moral alternativo al utilitarismo, teoría ética y económica, hasta entonces concebida como insuperable por un largo tiempo.

El objetivo principal de *Teoría de la Justicia* es desarrollar una teoría que pudiese servir como la base moral más apropiada para una sociedad democrática. La idea principal es que una sociedad justa tomará sus decisiones fundamentales en base al

respeto mutuo entre sus miembros. La propuesta de Rawls busca dilucidar lo que esta reciprocidad requiere a través de la formulación y defensa de principios fundamentales de justicia. Es claro que la inmensa influencia de esta obra en la filosofía política contemporánea es en gran parte explicada por el hecho de que logró un respeto filosófico generalizado para el proyecto de la justificación de principios morales.

El enorme peso del legado de Rawls es reconocido en este significativo párrafo de Robert Nozick, uno de sus más severos críticos,

Ahora los filósofos políticos tienen que trabajar según la teoría de Rawls, o bien, explicar por qué no lo hacen. Las consideraciones y distinciones que hemos desarrollado se esclarecen por la presentación magistral de Rawls de una concepción diferente y ayudan a esclarecerla. Incluso los que aún no están convencidos, después de luchar con la visión sistemática de Rawls, aprenderán mucho de un estudio cuidadoso. No hablo sólo del refinamiento...de los puntos de vista propios al combatir (lo que uno considera es) un error. Es imposible leer el libro de Rawls sin incorporar mucho, tal vez transformado, a la concepción profundizada de uno. Asimismo, es imposible terminar su libro sin una visión nueva y sugestiva de lo que la teoría moral puede intentar hacer y unir, sin una visión de lo *hermosa* que puede ser una teoría completa.<sup>125</sup>

---

<sup>125</sup> Robert Nozick, *Anarquía, Estado y Utopía*. Tr. Rolando Tamayo. México: FCE, 1991, p.183.

## BIBLIOGRAFÍA

- Arrow, Ken. "Rawls's Principle of Just Saving", en *Institute for Mathematical Studies in the Social Sciences, Stanford University, Technical Report no. 106*, Sep. 1973a.
- Barry, Brian. *The Liberal Theory of Justice, A Critical Examination of the Principal Doctrines in 'A Theory of Justice' by John Rawls*. Londres: Oxford University Press, 1973.
- Binmore, Ken. *Game Theory and the Social Contract, Vol I: Playing Fair*. Boston: Massachusetts Institute of Technology, 1994.
- . *Natural Justice*. EUA: Oxford University Press, 2005.
- Grcic, Joseph. "Rawls and Hobbes on Political Power," <<https://periodicos.ufsc.br/index.php/ethic/article/viewFile/17444/16066>>.
- Gruoso, Delfín. *Rawls, Legado de un Pensamiento*. Cali: Ed. Universidad del Valle, 2005.
- Habermas, Jürgen. "Reconciliación Mediante el Uso Público de la Razón," en *Debate Sobre el Liberalismo Político*, traducido por Gerard Vilar Roca. Barcelona: Paidós, 1998.
- Harsanyi, John. "Can the maximin principle serve as a basis for morality? A critique of John Rawls' theory," en *American Political Science Review*, 59, 1975.
- . *Essays on Ethics, Social Behavior, and Scientific Explanation*. Holland: D. Reidel Publishing Company, 1980.
- . "Bayesian Decision Theory, Subjective and Objective Probabilities, and Acceptance of Empirical Hypotheses," en *Synthese* 57, 1983.
- Hobbes, Thomas. *Leviatán*, traducido por Manuel Sánchez Sarto. México: FCE, 2002.
- Locke, John. *Dos Tractos Sobre el Gobierno*, traducido por Daniel Mielgo. Madrid: Biblioteca Nueva, 2015.

- Moehler, Michael. "The Rawls–Harsanyi Dispute: A Moral Point of View," en *Pacific Philosophical Quarterly*, doi: 10.1111/papq.12140. 2015.
- Nozick, Robert. *Anarquía, Estado y Utopía*, traducido por Rolando Tamayo. México: FCE, 1991.
- Paricio, Javier y Fernández Barreiro, Alejandro. *Historia del Derecho Romano y su Recepción Europea*. Madrid: El Faro, 2002.
- Rawls, John. *Teoría de la Justicia*, traducido por María Dolores González. México: FCE, 1978.
- . *El Liberalismo Político*, traducido por Vicente Orellana Aranda. Barcelona: Crítica, 2004.
- . *Political Liberalism*, expanded edition. New York: Columbia University Press, 2005.
- . *La Justicia como Equidad: Una Reformulación*, traducido por Andrés de Francisco. Barcelona: Paidós, 2012.
- . "Concepts of Distributional Equity: Some Reasons for the Maximin Criterion." *The American Economic Review*, vol. 64, no. 2, Papers and Proceedings of the Eighty-sixth Annual Meeting of the American Economic Association, May, 1974.
- . "Outline of a Decision Procedure for Ethics," en *Philosophical Review*. April 1951, 60 (2), p. 177-197.
- Rawls, John & M. McMurrin, Sterling. *Liberty, Equality and Law: Selected Tanner Lectures on Moral Philosophy*. Salt Lake City: University of Utah Press, 1987.
- Real Academia Española, *Diccionario de la Lengua Española*. Madrid: Espasa, 2014.
- Resnik, Michael. *Choices: An Introduction to Decision Theory*. EUA: University of Minnesota Press, 1987.
- Rodríguez, Jesús. *El Igualitarismo Liberal de John Rawls*. México: UAM, 2010.
- Rousseau, Jean-Jacques. *Discurso sobre el Origen y los Fundamentos de la Desigualdad entre los Hombres*, traducido por Ángel Pumarega. Valencia: Tilde, 1997.
- . *Del Contrato Social*, traducido por Mauro Armíño. Madrid: Alianza, 2012.

- . *Emilio, o de la Educación*, traducido por Mauro Armiño. Madrid: Alianza, 2005.
- Sandel, Michael. *El Liberalismo y los Límites de la Justicia*, traducido por María Luz Melón. Barcelona: Gidesa, 2009.
- Sarkar, Sumit y Chakraborty, Soumyakanti. "Does Rawls' 'Original Position' Induce Fairness? Experimental Findings on Selection Criteria in a Discrete Nash Demand Game Played from Behind the 'Veil of Ignorance'." (March 18, 2016). DDRN:<http://ssrn.com/abstract=2749496>.
- Schrader-Frechette, Kristine. *Risk and Rationality, Philosophical Foundations for Populist Reforms*. Los Angeles: University of California Press, 1991.
- Van Parijs, Philippe. "Difference Principles," en *Oxford Companion to Rawls*, editado por Samuel Freeman. New York: Cambridge University Press, 2003.