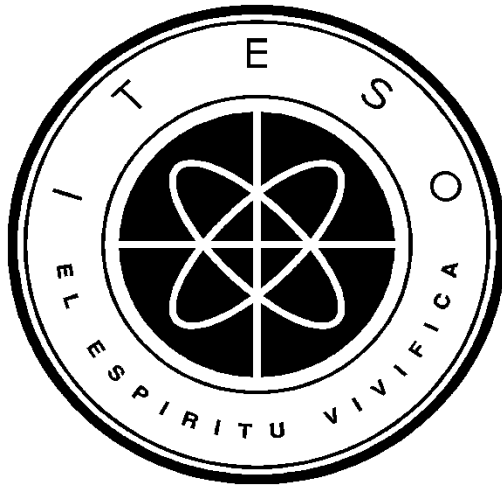


**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS
SUPERIORES DE OCCIDENTE**

RECONOCIMIENTO DE VALIDEZ OFICIAL, ACUERDO S.E.P. NO.
15018 PUBLICADO EN EL DIARIO OFICIAL DE LA FEDERACION
EL 29 DE NOVIEMBRE DE 1976.



DEPARTAMENTO DE FILOSOFÍA Y HUMANIDADES
MAESTRÍA EN FILOSOFÍA Y CIENCIAS SOCIALES

John Rawls *versus* utilitarismo:
La justicia examinada desde la “posición original”

TESIS PROFESIONAL PARA OBTENER EL GRADO DE
MAESTRO EN FILOSOFÍA Y CIENCIAS SOCIALES

Presenta:

Jaider Javier Salas Restrepo

Tlaquepaque, Jalisco, octubre de 2022

Índice

Introducción	5
Capítulo 1. El utilitarismo clásico frente a la teoría de la justicia de John Rawls	8
1. Jeremy Bentham: el placer y el dolor como directrices	8
2. John Stuart Mill: un redireccionamiento del utilitarismo	11
I. La justicia y la utilidad en Mill	14
3. Principio de utilidad según Henry Sidgwick	16
4. Críticas de Rawls a la doctrina utilitarista clásica	20
I. Principales parámetros de la concepción política de la justicia de Rawls	24
II. Bienes sociales primarios: una afrenta al utilitarismo	28
Capítulo 2. La posición original utilitarista desde el planteamiento ético de John Harsanyi	31
1. John Harsanyi: el utilitarismo de la preferencia	31
I. Modelo de equiprobabilidad	34
II. Comparaciones interpersonales de utilidad y el postulado de similitudes en Harsanyi	39
2. Principales críticas al utilitarismo contemporáneo: Gauthier frente a Harsanyi y Rawls	43
Capítulo 3. Configuración de la PO rawlsiana: imposibilidad de una sociedad imparcialmente justa desde una PO utilitarista	50
1. Posición Original rawlsiana	51
I. Velo de la ignorancia: dispositivo de imparcialidad	58
2. Principios de justicia: una alternativa a la propuesta utilitarista	60
I. La regla maximin como directriz: primera comparación fundamental	63
II. Principio de diferencia <i>versus</i> principio de utilidad restringida: segunda comparación fundamental	69
3. Crítica y contraargumentación de Harsanyi a la PO de Rawls	75
Conclusiones	84
Fuentes documentales	93

Introducción

La sociedad es un proyecto conjunto que vamos construyendo día a día. Es una casa común donde se encuentran distintas formas de concebir la vida. Es el escenario de batallas eternas por la defensa de valores que nunca pasarán de moda: el amor, la paz, la compasión, la igualdad, la justicia, entre muchos otros. Dentro de la sociedad tenemos la posibilidad de diseñar la mejor versión posible de nosotros mismos y, de esta manera, disfrutar del trozo de tiempo que nos ha sido dado para vivir. Mi planteamiento de investigación parte del deseo de anudar esfuerzos a los cambios posibles, a los horizontes imaginables en una democracia que demanda nuevos modos de organización en aspectos esenciales de su forma establecida.

Algunos de los rasgos distintivos de la manera en que nos hemos constituido en la actualidad como sociedad son el afán de tener, de consumir, de poseer, de sentir que estamos viviendo una vida buena y, por ende, feliz. Sin embargo, también podemos observar que otros de los rasgos salientes son las inevitables consecuencias de la carrera despiadada de cada individuo por obtener los mayores premios en los afanes mencionados. Como parte de tales consecuencias encontramos la creciente brecha de desigualdad social. También observamos violaciones a derechos y libertades básicos, arrodillados a intereses mezquinos de individuos que ostentan algún tipo de poder institucional –político, social o económico. De igual forma, nos encontramos situados en sociedades democráticas donde las expectativas de muchas personas se frustran ante la fría falta de oportunidades para soñar una vida mejor. Éstas son solo algunas de las características del rostro que presentan nuestras actuales democracias.

En el tiempo que llevo estudiando filosofía, uno de los temas que ha captado mi atención ha sido la propuesta de una teoría de la justicia del filósofo estadounidense John Rawls, quien construyó su edificio argumentativo en el marco de una sociedad democrática. Específicamente, me llamó la atención la manera en que, a partir de un diagnóstico profundo de la sociedad en la cual vivía, Rawls sintió la necesidad de elaborar una propuesta ética y política frente al paradigma utilitarista, predominante en su momento. El filósofo de Baltimore emprendió un camino de lucha por la justicia social valiéndose de un arma tan contundente como las ideas. Decidió defender una sociedad imparcialmente justa dándole un papel central a las instituciones políticas y sociales.

A lo largo de toda su teoría, Rawls trata de establecer una base consistente para abogar por la posibilidad de una sociedad realmente justa y equitativa. Para ello, el filósofo estadounidense acude a un mecanismo de representación denominado “posición original”, donde los representantes ideales (ficticios) de la sociedad deberán elegir los mejores principios de justicia para una sociedad democrática. En la posición original rawlsiana se escogerán, entonces, principios de justicia que coadyuvarán a establecer términos para la configuración de una sociedad imparcialmente justa.

A juicio de Rawls, una sociedad orientada por criterios utilitaristas sólo lograría crear más injusticias e iniquidades, por lo cual él ve la imperiosa necesidad de darle un rostro alternativo a la justicia con sus propios argumentos. La crítica fundamental de Rawls al utilitarismo remite a las principales posturas de su tradición clásica. Sin embargo, después de que el filósofo de Harvard hubiese establecido su teoría de la justicia, surgieron, a su vez, otras posturas utilitaristas que trataban de hacer frente a las críticas dispensadas por el filósofo de Baltimore al utilitarismo clásico.

Uno de los principales autores que confrontó las críticas de Rawls fue el economista húngaro John Harsanyi, el cual centra su discurso inspirado en aportes tanto del campo económico como del filosófico, buscando fundamentar una versión de sociedad donde las ideas utilitaristas podrían cumplir un papel fundamental en el terreno de la justicia social. Rawls respondió, en algunas ocasiones, a las críticas de Harsanyi; no obstante, es una discusión que aún sigue vigente y que no es ajena a los principales problemas que se cuestiona la filosofía moral y política en la actualidad.

A partir de este contraste descrito nacen mis deseos de profundizar en la crítica de Rawls al utilitarismo, así como de analizar la manera en que este último responde con la propuesta de un tipo de sociedad que, según sus argumentos, será más justa que la rawlsiana. Este análisis tendrá como escenario central la “posición original” en tanto mecanismo de representación que, según Rawls, ayudará a establecer los términos equitativos de la cooperación en una sociedad bien ordenada. De esta forma, tendremos como eje transversal en nuestra argumentación la manera en que se establece la posición original desde el constructo teórico de Rawls, pero también desde la propuesta de Harsanyi, para, de este

modo, contar con las suficientes herramientas argumentativas para hablar de la posibilidad de una sociedad justa bajo parámetros utilitaristas contemporáneos.

Por consiguiente, en este trabajo quisiera evaluar el alcance posible de una sociedad justa bajo los criterios de una posición original utilitarista. Para lograr tal objetivo, en el primer capítulo explicaré las principales ideas de la doctrina utilitarista clásica. Es un ejercicio que considero necesario como base del análisis de la crítica de Rawls a esta doctrina y como precedente de comprensión del utilitarismo contemporáneo. En el segundo capítulo expondré los principales argumentos de este último desde los aportes de Harsanyi; esta exposición tiene la intención de ubicarnos y preparar la entrada al análisis de una posición original utilitarista. En el tercer capítulo, finalmente, profundizaré en los puntos cruciales de la discusión entre el utilitarismo de Harsanyi y la teoría de la justicia como imparcialidad. De esta manera, en las conclusiones argumentaré por qué no podríamos hablar de una sociedad justa bajo una posición original utilitarista mediante la valoración de las críticas recíprocas de cada autor y algunos matices problemáticos que se desprenden de ella.

En ocasiones podríamos pensar que aportes como los de John Rawls han sido superados por la ingente literatura y las muchas discusiones que ha suscitado su pensamiento. Algo que no podemos negar, empero, es el hecho de que sus ideas aún no pierden vigencia. Una de las principales razones de esta vigencia tiene que ver con los profundos contrastes sociales, económicos y políticos que vivimos como democracias. Específicamente, por un lado, la eterna batalla existente entre individuos que sueñan con una sociedad donde se tenga como principal paradigma la maximización del bienestar tanto individual como general, dejando de lado – y muchas veces instrumentalizando- a un gran número de seres humanos que apenas luchan diariamente para sobrevivir; y, por otro lado, individuos que sueñan con una sociedad donde las principales decisiones a nivel institucional tengan como referencia a todos aquellos y aquellas que comúnmente son llamados “pobres”. Es por lo que este trabajo analiza, con detenimiento, cada una de las posiciones en el contraste descrito, buscando establecer una postura razonable en orden del fortalecimiento de una sociedad que desea caminar bajo las banderas de la justicia y la equidad social.

Capítulo 1. El utilitarismo clásico frente a la teoría de la justicia de John Rawls

En este capítulo nos proponemos explicar, en términos generales, las principales posturas alrededor de lo que se conoce como “utilitarismo”. Veremos, entonces, de qué manera se estableció la doctrina utilitarista como uno de los principales temas de discusión en filosofía a partir de finales del siglo XVIII. Para ello, abordaremos algunos de sus principales exponentes; a saber: Jeremy Bentham, John Stuart Mill y Henry Sidgwick. De esta manera, tendremos las herramientas suficientes para analizar la crítica que hace John Rawls a la doctrina clásica utilitarista e, igualmente, podremos ir estableciendo las bases suficientes para la disputa rawlsiana con la teoría utilitarista contemporánea.¹

1. Jeremy Bentham: el placer y el dolor como directrices

Cuando se habla de utilitarismo uno de los primeros nombres que sobresalen es el de Jeremy Bentham. Aunque este filósofo, economista y escritor inglés, nacido en Londres en 1748, no fuese el primero en hablar del principio de utilidad, se propuso, no obstante, exponer cómo tal principio podría repercutir en las disciplinas jurídicas. En una de sus principales obras, *An introduction to the Principles of morals and legislation*, de 1789, Bentham retomó grandes esfuerzos de argumentación que ya habían realizado otros filósofos alrededor de lo que luego se conocería como utilitarismo² para, a partir de allí, formular su postura.

Como se ha señalado, sería insensato atribuirle la doctrina de la utilidad general en tanto fundamento de la moralidad a Bentham. La utilidad ha sido un tema que, de una u otra forma, han abordado muchos filósofos. Sin embargo, como lo afirma Mill, “ha sido Bentham el que por vez primera introdujo precisión de pensamiento en la filosofía moral y política”.³

¹ Es oportuno mencionar que no haremos una revisión sistemática de cada autor utilitarista mencionado. Los hemos escogido porque coadyuvaron en la construcción de los principales argumentos que definen a la doctrina utilitarista clásica. Por lo tanto, haremos una exposición, *grosso modo*, de sus principales argumentos acerca de la manera en que pensaron el principio de utilidad.

² John Stuart Mill menciona, en su libro *Bentham*, que la idea del principio de utilidad general la tomó Bentham de Helvetius (*Cfr.*, John Stuart Mill, *Bentham*, Estudio preliminar, traducción y notas: Carlos Mellizo, Tecnos, 1993, p. 27). Además, señala que antes de Bentham existía suficiente literatura concerniente a la utilidad. Existían, por ejemplo, los *Ensayos sobre las características* de Brown, en el que se defendía arduamente la utilidad. Incluso, antes de Epicuro se podrían rastrear opiniones relativas a este tipo de posturas (*Ibidem*, p. 27). Nuestra intención no es profundizar en aquellas doctrinas antiquísimas; antes bien, nos proponemos abordar la discusión desde los parámetros que se establecen en la modernidad con los autores que hemos mencionado.

³ *Ibidem*, p. 28.

Por ende, detengámonos ahora en la manera en que el filósofo y economista inglés planteó el principio de utilidad. Habría que empezar diciendo que para Bentham, el hombre no es jamás entendido como un ser capaz de perseguir, como fin último, la perfección espiritual,⁴ sino como un ser susceptible de placeres y dolores, gobernado por el propio interés y las pasiones –que, a su vez, son egoístas-, así como por simpatías o antipatías hacia otros seres.⁵

Para Bentham, entonces, el dolor y el placer han sido puestos en el hombre y estos, a su vez, nos deberían llevar a establecer la normatividad acerca de lo que es bueno y lo que no lo es. El filósofo inglés sustenta esta visión retomando la argumentación de Locke acerca de la naturaleza de las ideas. Su razonamiento consiste en que, partiendo de la premisa de la existencia de las ideas simples y las ideas complejas, sabemos que las ideas simples no pueden ser descompuestas y que las ideas complejas están conformadas por la asociación de ideas simples. Las ideas simples son sensibles, pero también existen las que representan cualidades primarias de los objetos y otras que expresan solo cualidades de la mente o cualidades secundarias.⁶ Ahora, siendo el bien y el mal ideas complejas, ¿de qué ideas simples pudieron haber surgido? Bentham responde que surgieron de las ideas simples de placer –o una sensación agradable– y dolor –o una sensación desagradable–. Por lo tanto, el dolor y el placer serían modificaciones de la mente y no cualidades propias de los objetos.⁷

Las ideas de Bentham no tuvieron mucha acogida en su momento por algunos sectores de la sociedad, especialmente por el republicanismo clásico. Este último no suscribía la idea de que el hombre estuviese gobernado por el placer y el dolor, ya que “una vida bajo el imperio del placer y del dolor es propia de animales o de los esclavos, ya sea porque sus almas no poseen funciones inteligentes o porque no tienen o no fueron educados para tenerlas en grado suficientemente desarrollado como para dirigir el alma”.⁸ Para el republicanismo clásico, por el contrario, la felicidad proviene de la conquista de un bien supremo, que constituye algo diferente al solo placer o dolor.

⁴ Cfr., *Ibidem*, p. 46.

⁵ *Ibidem*, pp. 45-46.

⁶ Cfr., John Locke, *Ensayo sobre el entendimiento humano*, Fondo de Cultura Económica, México, 1956, pp. 110 – 122.

⁷ Cfr., Cícero Araújo, “Bentham: el utilitarismo y la filosofía política moderna” en *La filosofía política moderna. De Hobbes a Marx*, CLACSO, Consejo Latinoamericano de Ciencias Sociales, Buenos Aires, 2000, Tomado de <http://bibliotecavirtual.clacso.org.ar/clacso/se/20100609023007/11cap10.pdf>, p. 272.

⁸ *Ibidem*, p. 273

Bentham se opuso a ese tipo de críticas, ya que descarta la premisa de que un bien supremo sea la justificación para distinguir a la felicidad del placer. Además, se distancia de los republicanos clásicos en el concepto de “comunidad política”. Estos últimos pensaban, desde Aristóteles, que el ciudadano está individualmente subordinado al cuerpo político, lo que equivale a decir que su felicidad depende de la felicidad común y resulta inseparable de ésta.⁹ La felicidad es el resultado de una actividad del alma dirigida por su parte intelectual, y las funciones intelectivas del alma corresponden a las actividades deliberativas de la polis;¹⁰ por ende, será solo a través de la polis que el hombre podrá tener una vida buena. Bentham, por su parte, reconsidera el concepto de comunidad tal como lo acabamos de describir. Para él, entonces, la comunidad no es un cuerpo con alma, el cual piensa y siente, sino que está constituida por individuos, que son los que realmente piensan y sienten; son estos individuos los que, además, buscan el placer y tratan de huir a toda costa del dolor.¹¹

Por consiguiente, para el filósofo inglés “la felicidad de la comunidad no puede estar relacionada con un bien apartado de los individuos, sino que debe ser una simple suma de felicidades individuales... Cuanto mayor la suma, mayor la felicidad de la comunidad ”.¹² De esto habla el filósofo cuando se refiere al cálculo de felicidad: al hecho de pensar en la felicidad de cada uno de los individuos de la comunidad política y en cómo la suma de estas felicidades individuales redunda en la felicidad de todos.

Bentham se abstiene de proponer una definición del bien a la manera, por ejemplo, de Aristóteles. Es decir, Bentham no cree que se pueda plantear solo un camino para buscar la felicidad de todos. Algunos ideales funcionan para unos y no para otros: “no existe un estilo de vida colectivo que, como en la visión de inspiración aristotélica, corresponda a la perfección de la especie”.¹³ La legislación, por ejemplo, no debe ser un fin en sí mismo. Antes bien, debe ser un medio para la obtención de la mayor felicidad de la comunidad política. En este sentido, al momento de emitir leyes, la idea es que los beneficios de esa ley sean mayores que los costos que representarán para la comunidad; en este caso, de qué manera los impuestos destinados a estas leyes ayudarían al incremento del beneficio de la comunidad y

⁹ *Cfr., Ibidem*, p. 274.

¹⁰ *Cfr., Ibidem*, p. 275.

¹¹ *Cfr., Idem*.

¹² *Idem*.

¹³ *Ibidem*, p. 276.

no lo contrario. En efecto, vemos cómo "la moral utilitarista es, sí, teleológica, pero por consecuencialista y no por perfeccionista (como la moral republicana clásica)".¹⁴

Bentham piensa, entonces, que el placer conduce a la obtención de la felicidad. Sin embargo, ante la imposibilidad de que todos los individuos sean felices es preferible, por tanto, la felicidad del mayor número a la del menor número de individuos. Para ello, el Estado ha de ser el principal responsable de la promoción de la felicidad para la mayor cantidad de ciudadanos, y esto lo hará legislando bajo la inspiración del principio utilitarista. El filósofo inglés propone resaltar la importancia del interés común tanto de los que gobiernan como de los gobernados, y considera que el mejor sistema que ayudaría a tal objeto sería la democracia representativa. Otros tipos de configuraciones estatales, tales como un gobierno aristocrático o un gobierno mixto, velarían solo por el interés de una minoría dejando de lado a una gran parte de la sociedad y sus intereses comunes. Estas ideas las confrontaría después John Stuart Mill centrándose, fundamentalmente, en la manera en que Bentham propone el principio utilitarista. Veamos, entonces, los principales postulados de Mill.

2. John Stuart Mill: un redireccionamiento del utilitarismo

John Stuart Mill (J.S. Mill), nacido en 1806, heredó de su padre, James Mill, el conocimiento de la doctrina utilitarista; este último había sido seguidor de las ideas de Bentham. J.S. Mill, a su vez, introduciría algunos cambios en lo que propone Bentham acerca de la doctrina utilitarista. Nos detendremos, ahora, en las principales críticas de Mill a Bentham y en los principales argumentos de su propuesta en el marco de la doctrina utilitarista.

En *El utilitarismo*¹⁵, J.S. Mill expresa que desde Epicuro hasta Bentham se ha relacionado lo útil con el placer y con un necesario escape del dolor.¹⁶ Para este tipo de planteamientos, lo correcto e incorrecto tiene que ver con todo aquello que conduce a promover la felicidad y lo que no, respectivamente. Esta es la idea esencial del utilitarismo que enseñó Bentham. Es a partir de este escenario que J.S. Mill formulará algunas críticas al utilitarismo benthamiano y propondrá su propio concepto de utilidad.

¹⁴ *Idem.*

¹⁵ John Stuart Mill, *El utilitarismo*, Alianza Editorial, Madrid, 1984.

¹⁶ *Ibidem*, p. 45.

Mill acoge el principio de utilidad tal como lo plantea Bentham; no obstante, ve la necesidad de algunos cambios importantes en la forma en que Bentham lo dio conocer en su marco teórico. Lo primero que problematiza Mill es el hecho de que se ubique a la utilidad como único principio rector de la acción moral del hombre, idea que dio sustento a todo el edificio argumentativo benthamiano. Concretamente, ante el acento de Bentham en las consecuencias de las acciones en tanto consideración determinante de la moralidad de éstas, ¹⁷ Mill arguye la probabilidad de que Bentham se haya encerrado en los límites de las lógicas de su propio mundo y le haya faltado “un mayor conocimiento de la formación del carácter y de las consecuencias de las acciones sobre el estado de ánimo del agente”.¹⁸ Reducir la utilidad a un “principio de consecuencias específicas”, como lo denomina Mill, quizá tendría éxito aplicado a la jurisprudencia y a la legislación; sin embargo, él arguye que el principio de utilidad debe ayudar a organizar las instituciones sociales básicas para que los miembros de la sociedad puedan adquirir un carácter tal que los lleve a practicar un comportamiento deseado.¹⁹

Otro de los errores de Bentham, según Mill, fue quedarse confinado en uno de los tres aspectos de la acción humana, esto es, en el aspecto moral.²⁰ Al parecer, para Bentham la norma moral era el faro que debía guiar cualquier tipo de reflexión con relación al principio de utilidad. Al dejar de lado el aspecto estético y simpático de la acción humana, Bentham, por ejemplo, equiparaba la poesía a una igual cantidad de placer en un juego infantil.²¹ Mill lo expresa de la siguiente manera:

(Bentham) escribía y sentía como si la norma moral no solamente debiera ser la más importante (que debería), sino la única; como si ella fuese la sola maestra de todas nuestras acciones y hasta de todos nuestros sentimientos; como si admirar o apreciar, despreciar o rechazar a una persona por una acción que no produjese un bien o un daño, o que no produjese un bien o un daño proporcionados al sentimiento albergado, fuese una injusticia y un prejuicio.²²

¹⁷ *Cfr.*, John Stuart Mill, *Bentham, Op.cit.*, p. 84.

¹⁸ *Idem.*

¹⁹ John Rawls, *Lecciones sobre la historia de la filosofía política*, Paidós, Barcelona, 2009, p. 318.

²⁰ Mill señala que “toda acción humana tiene tres aspectos: su aspecto moral, que se refiere a su bondad o maldad; su aspecto estético, que se refiere a la belleza; su aspecto simpático, que se refiere a sus cualidades amables. El primero apela de suyo a nuestra razón y conciencia; el segundo, a nuestra imaginación; el tercero, a nuestro sentimiento humanitario hacia el prójimo” (John Stuart Mill, *Bentham, Op. cit.*, p. 85). Estos tres aspectos serán importantes para entender la propuesta utilitarista de Mill.

²¹ *Ibidem*, p. 87.

²² *Ibidem*, p. 86

Para comprender lo que significa el principio de utilidad en Mill, revisemos su análisis sobre los placeres. Se pregunta, específicamente, acerca de la diferencia entre cantidad y calidad de los placeres. Para ello, va más allá de los placeres como meras sensaciones y propone analizarlos a partir de la fuente que los produce. De aquí que hable de dos tipos de facultades en tanto fuentes de los placeres: las superiores y las inferiores. Las primeras se refieren al intelecto, al sentir, a la imaginación y a los sentimientos morales; las segundas se relacionan con las necesidades y exigencias físicas, las cuales ayudan a generar los placeres sensitivos.²³ Bentham diría que buscar la fuente de un placer no tiene importancia; si la intensidad y la duración de dos placeres es igual, estamos hablando, entonces, del mismo placer.

Ahora, ¿un placer es mejor por su cantidad o por su calidad? ¿De qué manera se podrían diferenciar? Mill, al enfrentarse a este tipo de cuestionamientos, llega a la conclusión de que “entre dos placeres, si hay uno al que todos, o casi todos los que han experimentado ambos, conceden una decidida preferencia, independientemente de todo sentimiento de obligación moral para preferirlo, ese es el placer más deseable”;²⁴ en este sentido, “está justificado que asignemos al goce preferido una superioridad de calidad que exceda de tal modo al valor de la cantidad como para que ésta sea, en comparación, de muy poca importancia”.²⁵ Por ejemplo, si a un individuo le produce más placer la vida ascética que una vida de ocio, sabiendo que la primera conlleva mayores sufrimientos, podríamos concluir que la vida ascética tendrá preferencia para ese individuo, sin importar que a otros individuos les genere más placer una vida dedicada al ocio. De esta forma, cabe deducir que, para Mill, la felicidad no solo se reduce a meras sensaciones de placer sino, también, al tipo de placer por el cual opta un individuo.

El utilitarismo de Mill está fundado en un hedonismo universalista. Es decir, “la felicidad que constituye el criterio utilitarista de lo que es correcto en una conducta no es la propia felicidad del agente, sino la de todos los afectados”.²⁶ En efecto, desde el principio de utilidad de Mill, tanto las leyes y organizaciones sociales como la educación y la opinión pública deben ayudar a que el individuo no solo piense en lógica de su propia felicidad, sino

²³ John Rawls, *Lecciones sobre la historia de la filosofía política*, *Op.cit.*, p. 322.

²⁴ *Ibidem*, pp. 48-49.

²⁵ *Ibidem*, p. 49.

²⁶ *Ibidem*, p. 62.

también en la de los demás; ambas deben ir necesariamente unidas.²⁷ En este sentido, para Mill la justicia es un derivado del principio de utilidad.

I. La justicia y la utilidad en Mill

Mill critica el hecho de que algunos pensadores tiendan a absolutizar la justicia en tanto algo natural, inherente al hombre: "el sentimiento de justicia bien pudiera ser un instinto peculiar y requerir, no obstante, al igual que nuestros demás instintos, el ser controlado e iluminado por una razón superior".²⁸ Parece evidente que el ser humano posee sentimientos naturales de justicia, sin embargo, hay que saber cómo proyectamos tal sentimiento como criterio último de la conducta, sin perder de vista la premisa de que ambos, tanto los sentimientos como la conducta, están íntimamente conectados. Por lo tanto, Mill se propone determinar "si la justicia o la injusticia de una acción es algo intrínsecamente peculiar y distinto de todas sus demás cualidades, o solamente una combinación de algunas de dichas cualidades, presentadas desde un ángulo especial".²⁹ ¿Es la justicia un sentimiento *sui generis* que no necesita de otros sentimientos o ser guiada por algún criterio en especial?³⁰

Mill encuentra, por consiguiente, un factor común en la forma en que es concebida la justicia; a saber: el hecho de que la noción de la justicia varía de acuerdo con las personas, adecuándose en cada caso a la noción que ellas tengan de la utilidad.³¹ Por ejemplo, "aquellos que consideran que la utilidad requiere diferencias de rango, no consideran injusto que la riqueza y los privilegios sociales se repartan desigualmente".³² También están los que piensan que el gobierno es necesario y, por ende, no ven injusticia alguna en la existencia de las desigualdades que se originan al otorgar a los magistrados poderes no concedidos al pueblo.³³ No obstante, Mill descubre que, históricamente, la idea de una prohibición legal es la idea generatriz de la noción de justicia, aunque experimente diversas transformaciones antes de que dicha noción, tal como se da en un estado avanzado de la sociedad, resulte completa.³⁴

²⁷ Cfr., *Ibidem*, p. 63.

²⁸ John Stuart Mill, *El utilitarismo*, *Op. cit.*, p. 101.

²⁹ *Idem*.

³⁰ Cfr., *Ibidem*, p. 102.

³¹ Cfr., *Ibidem*, p. 107.

³² *Idem*.

³³ Cfr., *Idem*.

³⁴ Cfr., *Ibidem*, p. 110.

Una de las conclusiones a las que llega Mill –en relación a su principal pregunta acerca de la necesidad, por parte de la justicia, de otros sentimientos o de algún otro criterio externo- es que el sentimiento de justicia se refiere al “sentimiento natural de resarcimiento o venganza, que el intelecto y la simpatía hace extensible a todos los perjuicios, es decir, a todos los daños que se nos causan a nosotros a través de, o en unión de, la sociedad en su conjunto”.³⁵ Este sentimiento natural nos llevaría a rechazar todo tipo de acciones que nos parezcan desagradables; por lo tanto, “cuando este sentimiento natural se moraliza mediante la incorporación de un sentimiento social, solo actúa en el sentido que viene determinado por el bien general, de tal modo que las personas justas rechazan los daños causados a la sociedad”.³⁶ Esta idea, parafraseando a Kant, nos podría decir que “debemos modelar nuestra conducta conforme a una norma que todos los seres racionales pudiesen aceptar con beneficio para sus intereses colectivos”.³⁷

Para Mill, el interés más importante en una sociedad es la seguridad y esta puede ser garantizada gracias, principalmente, a un derecho moral que le pertenece a cada individuo. En efecto, “tener derecho es tener algo cuya posesión ha de serme defendida por la sociedad”,³⁸ Mill asevera que, frente a las objeciones ante esta proposición y los cuestionamientos del porqué debería ser así, no podría ofrecer otra razón que la utilidad general.³⁹ El filósofo inglés, en efecto, ubica a la utilidad como “la *ratio* última de la moralidad, en virtud de la cual, precisamente, la justicia y sus preceptos adquieren la más vital importancia”.⁴⁰ Lo hace al vislumbrar la necesidad del reconocimiento de algún criterio distinto ante las muchas y diversas ideas que mantienen naciones e individuos acerca de la justicia, ya que “en la mente de un mismo individuo la justicia no es una norma, principio o máxima únicos, sino una multitud de ellos que no siempre coinciden en sus dictados”.⁴¹ De esta manera, para Mill:

La justicia es el nombre de ciertas clases de reglas morales que se refieren a las condiciones sociales del bienestar humano de forma más directa y son, por consiguiente, más absolutamente obligatorias que ningún otro tipo de reglas que orienten nuestra vida. De hecho, la idea que hemos averiguado que constituye la

³⁵ *Idem*, p. 115.

³⁶ *Idem*.

³⁷ *Ibidem*, p. 116.

³⁸ *Ibidem*, p. 118.

³⁹ *Cfr.*, *Idem*.

⁴⁰ *Ibidem*, p. 130.

⁴¹ *Ibidem*, p. 120.

esencia de la justicia, a saber, un derecho que posee un individuo implica y testimonia esta obligación más vinculante.⁴²

Una de las críticas más conocidas al utilitarismo de Mill es que pareciera implicar la disposición a sacrificar a un individuo, o a un grupo de individuos, en aras del beneficio o el bien colectivo. No obstante, como se ha constatado, esta es solo una de las opiniones acerca de la justicia que, según Mill, tienen los individuos de una sociedad.⁴³ Siendo una de las tantas acepciones de justicia, no se podría sostener por sí sola frente a otras posiciones acerca de la justicia.

Otros filósofos y economistas, contemporáneos a Mill, se adentraron en la tarea de reflexionar a partir del principio de utilidad. Uno de éstos es Henry Sidgwick que, junto a Bentham y Mill, sería considerado uno de los principales teóricos del utilitarismo moderno. Analicemos, entonces, los argumentos más importantes que Sidgwick aporta a la doctrina utilitarista.

3. Principio de utilidad según Henry Sidgwick

El filósofo inglés Henry Sidgwick, nacido en Skipton (Reino Unido) el 31 de mayo de 1838, es conocido por presentar una de las obras más completas acerca de la doctrina utilitarista. *The Methods of Ethics*⁴⁴ (1874) ha sido, para muchos filósofos, el mejor tratado de teoría moral que jamás se haya escrito y uno de los clásicos de la filosofía inglesa.⁴⁵ Para Rawls, por ejemplo, la obra mencionada “constituye la formulación más clara y accesible de la doctrina utilitarista clásica”.⁴⁶ Es también por esto que Rawls, como veremos, decidirá criticar la doctrina utilitarista clásica representada, principalmente, en la visión de Sidgwick. Rawls expresa esta preferencia de su crítica de la siguiente manera: “Sidgwick es más consciente que otros autores clásicos de las múltiples dificultades que presenta esta doctrina

⁴² *Ibidem*, p. 126.

⁴³ *Cfr.*, *Ibidem*, 121.

⁴⁴ Henry Sidgwick, *The methods of ethics*, Macmillan and Co., New York, 1907. Tomado de <http://www.gutenberg.org/files/46743/46743-h/46743-h.htm>, el 06 de febrero de 2020.

⁴⁵ Bart Schultz, *The happiness philosophers: the lives and works of the great utilitarians*, Princeton University Press, New Jersey, 2017, p. 442.

⁴⁶ John Rawls, *Lecciones sobre la historia de la filosofía política*, *Op. cit.*, p. 462.

(el utilitarismo), y trata de abordarlas de forma sistemática y exhaustiva, sin desviarse en ningún momento de la doctrina estricta, como hizo, por ejemplo, J.S. Mill”.⁴⁷

Para Sidgwick, entonces, “el fin moral último de la acción social e individual es la obtención de la mayor suma neta posible de felicidad de todos los seres con capacidad de sentir”.⁴⁸ A partir de esta premisa, Sidgwick tomó distancia de las posturas utilitaristas tanto de Bentham como de J. S. Mill. A diferencia de éstos, “su principal énfasis lo puso en la esfera de la ética personal, la cual Bentham y Mill habían, relativamente, descuidado”.⁴⁹ Además, consideró el problema metafísico de la libertad de la voluntad irrelevante en la discusión ética, arguyendo que era un error pensar que los individuos eran siempre, o generalmente, psicológicamente motivados a actuar solo por sus propios gustos o por un bien personal.⁵⁰

El principal análisis que hace Sidgwick del aporte de Mill es, sobre lo que aquél denomina, “hedonismo psicológico” (*Psychological Hedonism*) y un “hedonismo ético” (*Ethical Hedonism*).⁵¹ El primero, el hedonismo psicológico, se refiere al hecho de que cada hombre busca su propia felicidad; mientras que el hedonismo ético hace referencia al deber de cada hombre de buscar la felicidad general.⁵² La consideración de estos dos tipos de hedonismos llevó a Sidgwick a examinar, metodológicamente, la relación entre el interés y el deber u obligación. Este dilema ético, el conflicto entre mi propia felicidad y la felicidad general -entre utilitarismo y egoísmo racional-, permanecería a la sombra del desarrollo de *The Methods*, ante el cual el autor afirmaría que, a razón de este dualismo de la razón práctica, parece no existir ninguna solución objetiva.⁵³ De esta manera, Sidgwick se vio en la necesidad de acudir a autores como Whewell, Butler, Kant y Aristóteles en aras de buscar una defensa razonada del sacrificio (*self-sacrifice*) y su relación con el sentido común moral (*common-sense morality*).⁵⁴

⁴⁷ *Idem.*

⁴⁸ *Idem.*

⁴⁹ Bart Schultz, *The happiness philosophers, Op. cit.*, p. 454.

⁵⁰ *Idem.*

⁵¹ *Cfr.*, *The methods of ethics, Op. cit.*, pp. 39 – 54.

⁵² Bart Schultz, *The happiness philosophers, Op. cit.*, p. 455.

⁵³ *Cfr.*, *Lecciones sobre la historia de la filosofía política, Op. cit.*, 468.

⁵⁴ *Cfr.*, *The happiness philosophers, Op. cit.*, 457.

Sidgwick cree en una genuina y fundacional verdad ética; esto lo hace reconciliando la moralidad intuitiva y el utilitarismo.⁵⁵ De esta manera, su principal obra- antes mencionada- se propondrá “examinar todos los distintos métodos de obtención de convicciones morales razonadas sobre lo que se debe hacer y lo que no se debe hacer que se pueden encontrar -explícita o implícitamente- en la conciencia moral de la humanidad en general”.⁵⁶ Para ello, el autor inglés define un “método de ética” como un “procedimiento racional a través del cual decidimos lo que los seres humanos individuales deberían hacer o lo que sería correcto que hicieran, o lo que deberían tratar de realizar actuando libre y voluntariamente”.⁵⁷ Con tal método, Sidgwick está interesado, principalmente, en reconocer “qué conclusiones alcanzaremos racionalmente si partimos de ciertas premisas éticas, y con qué grado de certeza y precisión lo haremos”,⁵⁸ además, a pesar de que la principal intención del filósofo inglés no fue plantear una justificación del utilitarismo clásico a través de su obra magna, fue siempre un fuerte defensor del utilitarismo y, por ende, llegó a conclusiones como ésta: “el utilitarismo, desde un punto de vista neutral, supera los criterios de un método racional de la ética mucho mejor que cualquier forma de intuicionismo (y, por lo tanto, es superior al intuicionismo)”.⁵⁹

Una de las conclusiones a las que llega Sidgwick es que tanto el utilitarismo como el egoísmo racional satisfacen, igualmente bien, los criterios objetivos de un método racional de la ética;⁶⁰ teniendo en cuenta que unas de las principales características que un método racional de la ética debe tener son los siguientes:

Deben ser, al menos, tan ciertos como cualesquiera otros principios morales; deben tener una validez superior a la de otros principios; deben ser, realmente, evidentes y no derivar su validez de ningún otro principio (...); deben controlar, regular y sistematizar otros principios y criterios subordinados (y preceptos y creencias morales de nivel inferior), a fin de organizarlos hasta formar un sistema completo y armonioso libre de elementos arbitrarios”.⁶¹

⁵⁵ Cfr., *The methods of ethics*, *Op. cit.*, pp. 423-459.

⁵⁶ John Rawls, *Lecciones sobre la historia de la filosofía política*, *Op. cit.*, p. 466.

⁵⁷ *Idem.*

⁵⁸ *Ibidem*, p. 467.

⁵⁹ *Idem.*

⁶⁰ Cfr., *Ibidem*, p. 468.

⁶¹ *Ibidem*, p. 470.

Sidgwick se propone evaluar los principios de justicia que se encuentran en el sentido común de las personas, tal como lo haría J. S. Mill al tratar de encontrarle un significado al concepto de justicia.⁶² El filósofo de Skipton concluye, entonces, que el principio que funciona genuina, racional y objetivamente ante otros muchos principios es el principio de utilidad.⁶³

Por consiguiente, para Sidgwick el principio ético debe llevar a maximizar el balance neto de placer sobre dolor.⁶⁴ Se trata de una suma de la medición de placeres y dolores como cantidades objetivas dadas por la intensidad y la duración de éstos, que siempre van a ser igual a 1, con el balance neto de felicidad para cada individuo a lo largo de un período de tiempo relevante.⁶⁵ La idea, entonces, es maximizar, lo más que se pueda, la utilidad a lo largo de un período de tiempo determinado, teniendo en cuenta que los placeres y dolores, como cantidades objetivas que se miden, son independientes de “cualquier relación objetiva de los individuos entre sí que pueda ser un condicionante de esas experiencias, o de los fines de los deseos cuya satisfacción o insatisfacción provoque el placer o el dolor”.⁶⁶ En la suma de la felicidad se incluye a todos los individuos que estén afectados por las instituciones.

Para poder maximizar la suma lineal de utilidades, según Sidgwick es necesario realizar una comparación interpersonal de utilidad, donde “cada individuo puede estimar y comparar sus propios niveles de felicidad basándose en su propia introspección y memoria: los placeres y los dolores son aspectos de los que tenemos conocimientos directo en experiencias que juzgamos agradables o desagradables”.⁶⁷ Estas sensaciones o experiencias se considerarán el bien supremo, lo cual implica que la doctrina utilitarista clásica define la felicidad o el bien supremo independientemente de cualquier noción relacionada con la justicia, lo correcto o cualquier virtud moral; de aquí podemos colegir, nuevamente, el rasgo marcadamente consecuencialista de la doctrina utilitarista.⁶⁸

El modo de hacer las comparaciones interpersonales de utilidad radica en el hecho de tener tanto una medida cardinal de utilidad para cada individuo como un modo de concordar las medidas de utilidad de individuos distintos, es decir, “unas reglas de correspondencia que

⁶² Cfr., *The methods of ethics*, *Op. cit.*, pp. 264-294.

⁶³ Cfr., *Lecciones sobre la historia de la filosofía política*, *Op. cit.*, p. 471.

⁶⁴ Cfr., *Ibidem*, p. 474.

⁶⁵ Cfr., *Ibidem*, pp. 474-475.

⁶⁶ *Ibidem*, p. 475.

⁶⁷ *Idem*.

⁶⁸ Cfr., *Ibidem*, p. 484.

nos indiquen cómo comparar y ponderar los placeres de diferentes personas”.⁶⁹ Una medida cardinal individual de utilidad⁷⁰ hace referencia a una suerte de cálculo que hace cada individuo de su propia felicidad; esto implica, por ejemplo, estimar la intensidad y duración de los estados de conciencia -sensaciones- agradables o desagradables.⁷¹

Por lo visto, entonces, para la doctrina utilitarista clásica, como ha sido planteada por Sidgwick, el principal objetivo de todo individuo racional que camine en búsqueda de su bienestar y felicidad es coadyuvar en la maximización de todo aquello que conduzca a la generación de placer (o sensación agradable). En este sentido, lo justo, para un utilitarista, sería todo aquello que conlleve a la potencialización del bien, es decir, a la satisfacción de los deseos racionales, “entendiendo por éstos aquellos deseos que los individuos tendrían si sometieran sus deseos reales presentes a ciertas formas de evaluación racional”.⁷²

4. Críticas de Rawls a la doctrina utilitarista clásica

Veamos, ahora, cuáles son las principales críticas que hace Rawls a la doctrina utilitarista clásica y, de esta manera, analizar sus principales motivos para no considerar el utilitarismo, desde la propuesta de Bentham, J.S. Mill y Sidgwick, principalmente, como una opción razonable para la construcción de una sociedad justa. Al distinguir tales argumentos, abordaremos simultáneamente la propuesta de la justicia como imparcialidad de Rawls.

Por consiguiente, la premisa fundamental del utilitarismo que Rawls analizará es que “cuando las instituciones más importantes de la sociedad están dispuestas de tal modo que obtienen el mayor equilibrio neto de satisfacción distribuido entre todos los individuos pertenecientes a ella, entonces la sociedad está correctamente ordenada y es, por tanto, justa”.⁷³ Como lo habíamos visto con Sidgwick, un individuo racionalmente querrá obtener el mayor beneficio posible y, de esta manera, podrá llevar a cabo sus propios planes de vida.

⁶⁹ *Ibidem*, p. 489.

⁷⁰ Es importante tener en cuenta que la medida cardinal de utilidad individual difiere de la medida de utilidad de Von Neumann y Morgenstern, la cual está “basada en unas elecciones coherentes entre loterías alternativas, es decir, entre combinaciones diversas de alternativas ponderadas por su probabilidad” (*Ibidem*, p. 490). Explicaremos con más detalle esta medida de utilidad cuando presentemos la argumentación de las posturas utilitaristas contemporáneas.

⁷¹ *Cfr.*, *Ibidem*, p. 489.

⁷² *Ibidem*, p. 487.

⁷³ John Rawls, *Teoría de la justicia*, Fondo de Cultura Económica, México, 1997, p. 34.

El papel de la sociedad será, entonces, ayudar a promover la satisfacción de los sistemas de deseos de los muchos individuos que pertenecen a ella.⁷⁴ En efecto, en esta concepción utilitarista “el bien es definido independientemente de lo justo, y entonces lo justo es definido como aquello que maximiza el bien”.⁷⁵ Siendo una concepción consecuencialista, lo justo siempre está supeditado a lo que se define como lo bueno.

Rawls señala que existe un modo de pensar en la sociedad que lleva a considerar la opción utilitarista como la más viable. Es decir, una persona está actuando correctamente cuando desea y trabaja en la obtención del mayor beneficio posible y, de esta manera, en la promoción de sus fines racionales.⁷⁶ En este sentido, así como para un individuo será prioritario aumentar su propio bienestar, de la misma forma la sociedad tendrá que encontrar la mejor manera de promover el bienestar de los individuos que pertenecen a ella. Por lo tanto, “del mismo modo en que un individuo equilibra ganancias presentes y futuras con pérdidas presentes y futuras, de ese modo una sociedad puede equilibrar satisfacciones e insatisfacciones entre individuos diferentes”.⁷⁷ Una sociedad utilitarista, a través de sus instituciones, trabajará, entonces, en aras de la maximización del equilibrio neto de satisfacción.⁷⁸ Por lo tanto, la justicia social es concebida como “el principio de prudencia racional aplicado a una concepción colectiva del bienestar del grupo”.⁷⁹

Una teoría moral teleológica, como lo es el utilitarismo, conduciría, según Rawls, al establecimiento de una sociedad que piensa desde lógicas de maximización, y en la cual se parte de lo que se considera bueno. En efecto, si una sociedad considera a la excelencia humana como el mayor bien, entonces tendremos una sociedad perfeccionista; si lo es el placer, se vivirá bajo los parámetros hedonistas; si es la felicidad, el eudemonismo, entre otras.⁸⁰

La visión utilitarista parece ser muy atractiva, ya que a nadie le gustaría no tener más que menos, especialmente cuando nos referimos al bienestar personal y, en adición, al social. No obstante, Rawls se pregunta por la lógica de distribución de la suma de satisfacciones a

⁷⁴ *Cfr., Ibidem*, p. 35.

⁷⁵ *Ibidem*, p. 36.

⁷⁶ *Cfr., Ibidem*, p. 35.

⁷⁷ *Idem*.

⁷⁸ *Cfr., Idem*.

⁷⁹ *Idem*.

⁸⁰ *Cfr.*, p. 37.

partir de la premisa utilitarista; deduce que en una sociedad utilitarista no podría existir un patrón de distribución de satisfacciones entre los individuos, es decir, cada persona buscaría la manera de aumentar, lo más que se pueda, su estado de bienestar. La sociedad será la encargada de poner los medios para que cada quien cumpla sus satisfacciones al máximo. Por ejemplo, cada decisión tomada por el órgano legislativo al formular una ley, tendría como propósito el aumento del mayor equilibrio de satisfacción.

En una sociedad utilitarista, por ende, la cooperación social estará definida por los términos que lleven a la mayor suma de satisfacción de los deseos racionales de los individuos.⁸¹ El único criterio para cualquier tipo de distribución que se quiera realizar estará determinado por la obtención del mayor equilibrio de satisfacción. Además, el principio que guiará el modo de proceder utilitarista será el de la elección racional de cada ciudadano.⁸² Tal principio lleva, necesariamente, a utilizar la figura del “espectador imparcial”⁸³ como el “individuo perfectamente racional que se identifica y tiene la experiencia de los deseos de otros como si fuesen los propios”.⁸⁴ En esta visión, tal como lo expresa Rawls:

La naturaleza del legislador imparcial no es, por tanto, materialmente distinta de la del empresario que decide cómo maximizar su ganancia mediante la producción de esta o aquella mercancía, o de la del consumidor que decide cómo llevar al máximo su satisfacción mediante la compra de esta o aquella colección de bienes.⁸⁵

Otra de las críticas que hace Rawls a la postura utilitarista es que, si pensáramos una sociedad dirigida por tal criterio, se presentarían casos donde, por ejemplo, la pérdida de libertad de unos se vería justificada por la búsqueda de un bien mayor que será compartido por otros.⁸⁶ En este sentido, para el utilitario, los “preceptos de justicia del sentido común y las nociones del derecho natural no tiene sino un valor subordinado como reglas secundarias”.⁸⁷ Es decir, el utilitarismo no tiene como eje central la prioridad de la justicia; antes bien, considera este tipo de prioridades, desde el principio de utilidad, “como una ilusión socialmente útil”.⁸⁸

⁸¹ *Cfr.*, *Teoría de la justicia*, *Op. cit.*, p. 37.

⁸² *Cfr.*, *Ibidem*, p. 38.

⁸³ Argumento utilizado, principalmente, por Adam Smith como recurso de imparcialidad en los juicios morales.

⁸⁴ John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 38.

⁸⁵ *Idem*.

⁸⁶ *Cfr.*, *Ibidem*, p. 39.

⁸⁷ *Idem*.

⁸⁸ *Idem*.

Rawls considera, de igual manera, que es insuficiente el principio de elección racional tal como es planteado por la doctrina utilitarista. Por esto propondrá un método de elección a través de un contrato original, que explicaremos más adelante. De igual manera, Rawls ve como una dificultad el hecho de que el utilitarismo sea una teoría teleológica. Ante esto propondrá una teoría deontológica, “que no especifica el bien independientemente de la justicia, o no interpreta lo justo como maximización del bien”.⁸⁹ Profundizaremos en este aspecto también al momento de abordar la propuesta rawlsiana en su conjunto.

Otra crítica de Rawls a la doctrina utilitarista clásica tiene que ver con el valor otorgado a la satisfacción del deseo. Este valor será importante al momento de decidir lo que es justo:⁹⁰ al momento de “calcular el equilibrio mayor de satisfacción no importa, excepto indirectamente, de qué son los deseos”.⁹¹ De aquí se seguiría, por ejemplo, que si uno de estos deseos es la discriminación racial o la esclavitud, “la satisfacción de estos deseos debe ser sopesada en nuestras deliberaciones de acuerdo con su intensidad, o lo que sea, al igual que otros deseos”.⁹² Rawls, por su parte, propondrá que los deseos y aspiraciones se restrinjan desde el inicio por los principios de justicia que especifican los límites que los sistemas de fines de los hombres tienen que respetar.⁹³

Como hemos analizado, Rawls toma como sugerencia de su crítica las principales premisas de la doctrina utilitarista clásica representada, principalmente, en filósofos como Bentham, J.S. Mill y Sidgwick.⁹⁴ Ahora analizaremos, *grosso modo*, el tipo de sociedad que propondrá Rawls a partir las insuficiencias y vacíos en las principales ideas del utilitarismo clásico. De esta manera, obtendremos claridad de la distancia que toma Rawls respecto del utilitarismo, para luego analizar algunas posiciones utilitaristas contemporáneas y tratar de responder la pregunta principal de esta investigación.

⁸⁹ *Ibidem*, pp. 40-41.

⁹⁰ *Cfr.*, *Ibidem*, p. 41.

⁹¹ *Idem*.

⁹² *Idem*.

⁹³ *Cfr.*, *idem*.

⁹⁴ No nos detuvimos con detalle en las principales críticas rawlsianas al utilitarismo porque volveremos a ellas en el segundo capítulo, retomándolas desde la visión contemporánea. Quisimos presentar, no obstante, los principales puntos en discusión, los cuales dieron pie al nacimiento de la propuesta rawlsiana.

I. Principales parámetros de la concepción política de la justicia de Rawls

Al hablar de justicia en la contemporaneidad es difícil prescindir del importante aporte que hizo John Rawls a este respecto. En su más representativa obra, *Teoría de la justicia*⁹⁵ (1971), expone sus principales argumentos para poder hablar de una sociedad equitativamente justa. En obras posteriores, tales como *Liberalismo político* (1993) y *La justicia como equidad: una reformulación* (2001), delimita y reformula partes importantes de su propuesta para, de esta manera, hacerla más inteligible y afable ante las inclementes críticas de filósofos y economistas que vislumbraron el nacimiento de una gran obra que alimentaría la reflexión en la filosofía política de los últimos tiempos.

Antes de proceder a explicar los principales elementos de la teoría rawlsiana, teniendo en cuenta las principales críticas que dirige a la tradición utilitarista clásica, detengámonos en lo que significa, para Rawls, la justicia. Para ello, sería conveniente despejar una duda técnica que reluce al momento de encontrarnos con distintas traducciones de lo que Rawls llamó *Justice as Fairness*. Algunos autores lo traducen como “justicia como equidad”,⁹⁶ o “justicia como imparcialidad”⁹⁷ o, incluso, como “justicia imparcialmente equitativa”.⁹⁸ Nuestro propósito es, entonces, averiguar cuál de estas traducciones al castellano tiene más sentido desde el concepto de justicia que presenta Rawls a lo largo de su pensamiento. Esto nos llevará, en efecto, a contrastar los conceptos de equidad e imparcialidad.

Antes de Rawls, algunos autores habían hablado de la imparcialidad estrechamente relacionada con la justicia. Uno de esos autores es J.S. Mill. Tomaremos su concepto de imparcialidad⁹⁹ porque nos parece que puede dar mucha claridad al momento de confrontarlo con la propuesta rawlsiana. Lo primero que tiene claro Mill es que sería casi imposible equiparar justicia y parcialidad. Denomina parcial al hecho de “mostrar favoritismo o

⁹⁵ En el resto del presente escrito usaremos la traducción al castellano que hace María Dolores González de *A theory of justice*, en la editorial Fondo de Cultura Económica.

⁹⁶ Como es el caso de la traducción que Andrés de Francisco de *Justice as fairness: a restatement*, en la editorial Paidós. Cfr. John Rawls, *La justicia como equidad: una reformulación*, Paidós, Barcelona, 2000.

⁹⁷ Tal como traduce María Dolores González en Fondo de Cultura Económica, en la referencia arriba citada.

⁹⁸ Ver, por ejemplo, el artículo de Fernando Rodríguez, titulado *El imparcial*, donde -prácticamente- está poniendo al mismo nivel la imparcialidad y la equidad. Cfr. Fernando Rodríguez Genovés, *El imparcial*, El Catoblepas: revista crítica del presente, número 117, noviembre 2011, página 7. Tomado de <http://www.nodulo.org/ec/2011/n117p07.htm>

⁹⁹ J.S. Mill usa el concepto de “impartiality”. Cfr., John Stuart Mill, *Utilitarianism*, Global Grey ebooks, 2015. Tomado de <https://www.globalgreybooks.com/utilitarianism-ebook.html>, el 20 de diciembre de 2019, p. 44.

preferencia respecto a una persona en detrimento de otra en cuestiones en las que el favoritismo y la preferencia no tienen propiamente cabida”.¹⁰⁰ En este sentido, la imparcialidad es considerada, no como un deber en sí misma, sino como un instrumento para otro deber.¹⁰¹ Mill entiende la imparcialidad como una obligación debida de la justicia, lo que puede significar que “influyan en nosotros únicamente las consideraciones que se supone debieran influir en el caso particular que tratemos, resistiéndonos a dejarnos condicionar por motivos distintos que promuevan conductas distintas de las que tales consideraciones requerirían”.¹⁰²

Además, para J.S. Mill la imparcialidad está directamente relacionada con la igualdad: “todo el mundo mantiene que la igualdad es una exigencia de la justicia, excepto cuando consideran que razones de conveniencia requieren la desigualdad”.¹⁰³ En este punto es cuando Mill explica que lo justo dependerá del concepto de lo útil que tenga cada persona y, por lo tanto, sería muy difícil contradecir algunas de esas opiniones sobre la justicia si no se tiene un criterio establecido que pueda ayudar a dirimir este tipo de discusiones. No obstante, ya habíamos hablado de la forma en que J.S. Mill resuelve dicho inconveniente. Lo que nos interesa, por ahora, es hacer notar la manera en que, casualmente, en Mill se vislumbran algunas de las características que se desarrollarían en la teoría rawlsiana.

Analicemos ahora, brevemente, en qué consiste la equidad. Si buscamos definiciones de lo que significa “equidad” nos encontraremos con palabras comunes, tales como igualdad, respeto a las diferencias, el merecimiento, entre otras.¹⁰⁴ A simple vista, y siguiendo la traducción al castellano de algunos autores que han escrito sobre la teoría rawlsiana, es sensato colegir que imparcialidad y equidad están estrechamente unidas dentro del planteamiento de una sociedad justa que hace Rawls. El matiz de “imparcialidad” se encuentra reflejado en el mecanismo de elección de los principios de justicia que propone el filósofo estadounidense; estamos hablando de la “posición original” donde, a través del “velo de la ignorancia”, representantes ideales de la sociedad deberán elegir los principios de

¹⁰⁰ John Stuart Mill, *El utilitarismo*, *Op. cit.*, 106.

¹⁰¹ *Cfr.*, *Idem*.

¹⁰² *Ibidem*, p. 107.

¹⁰³ *Idem*.

¹⁰⁴ *Cfr.*, *Diccionario de la Real Academia Española (RAE)*, Tomado de <https://dle.rae.es/equidad>, el 06 de febrero de 2020.

justicia que crean más razonables y racionales para una sociedad bajo un pluralismo razonable.¹⁰⁵

Sin embargo, partiendo de este escenario de imparcialidad, otro de los rasgos característicos de la concepción política rawlsiana es la búsqueda de la igualdad. Esto lo hace partiendo de una distribución igualitaria de los “bienes sociales primarios”,¹⁰⁶ de manera que una redistribución de éstos redunde en un mayor beneficio para toda la sociedad, donde los menos aventajados no se queden rezagados en la distancia abismal que genera la desigualdad. Silvina Ribotta afirma que Rawls, con la posición original, busca “establecer un procedimiento equitativo- *Justice as Fairness*- mediante el cual los principios de justicia que se elijan sean siempre justos. Un escenario igualador que garantice principios de justicia equitativos”.¹⁰⁷ De esta manera, siguiendo nuestro razonamiento, nos parecería prudente, para ser más precisos y hacer justicia a la pretensión de Rawls, usar la traducción “justicia como imparcialidad”; no obstante, nos referimos al tipo de sociedad que sueña Rawls como una “sociedad equitativamente justa”, de esta manera haremos venia a la relación existente entre estos dos términos en el marco de la teoría rawlsiana.

Entonces Rawls, al equiparar la justicia a la imparcialidad, está queriendo decir que la justicia no puede ser usada en función de cualquier tipo de concepción de bien, tal como lo propuso el utilitarismo. La prioridad de lo bueno sobre lo justo podría llevar a caer en errores tales como los que mencionábamos; pensar, por ejemplo, que, ya que es bueno el mayor equilibrio de satisfacción en una sociedad, entonces, será justo privar de la libertad a unos a costa de otros muchos. La justicia es, en efecto, la primera virtud de las instituciones sociales, “como la verdad lo es de los sistemas de pensamiento (...) no importa que las leyes e instituciones estén ordenadas y sean eficientes: si son injustas han de ser reformadas o abolidas”.¹⁰⁸ De esta manera, sería injusto que los sacrificios impuestos a unos sean compensados por la mayor cantidad de ventajas disfrutadas por muchos.¹⁰⁹ Una sociedad justa es aquella que cuenta con unas instituciones que garanticen libertades básicas iguales a todos los ciudadanos pertenecientes a una sociedad democrática; además de otros principios

¹⁰⁵ Más adelante ahondaremos en estos conceptos.

¹⁰⁶ Más adelante profundizaremos en este concepto.

¹⁰⁷ Silvina Ribotta, *John Rawls: sobre la (des)igualdad y justicia*, Dykinson, Madrid, 2009, p. 59.

¹⁰⁸ John Rawls, *Teoría de la justicia*, *Op. cit.*, 17.

¹⁰⁹ *Cfr.*, *Idem*.

por los que las principales instituciones de la sociedad deben velar. Más adelante ahondaremos en estos principios de justicia.

Hay que tener en cuenta que Rawls se ubica dentro de los límites de una sociedad democrática, la cual se caracteriza, principalmente, por un pluralismo razonable; éste evitará que el control social lo tome cualquier forma de totalitarismo, despotismo o sistema dirigido por la defensa de una determinada doctrina comprensiva y que excluya los diferentes modos de concebir la vida. A partir del pluralismo razonable, por el contrario, una sociedad democrática está conformada de tal manera que puede albergar distintas concepciones de bien que posean los individuos estableciendo, en medio de las diferencias, acuerdos políticos de justicia que redunden en el beneficio de todos los ciudadanos.

Una sociedad democrática tampoco podría ser una asociación, ya que a esta última se entra voluntariamente, mientras que un individuo nace siendo parte de aquella. En este sentido, afirma Rawls, la filosofía política tiene el papel de reconciliar los sentimientos de insatisfacción o malestar frente a cómo se ha constituido la democracia a través de sus principales instituciones a lo largo del tiempo y, de esta manera, de recuperar la confianza en el papel que puedan llegar a tener las principales instituciones sociales al momento de pensar en la construcción de una sociedad justa.

A las principales instituciones políticas y sociales Rawls le denomina “estructura básica de la sociedad”. Esta estructura es parte del modo en que estas instituciones se acoplan para formar un sistema de cooperación, la manera en que ayudan en la asignación de derechos y deberes básicos, así como la forma en que regulan la división de las ventajas que surjan de la cooperación social a lo largo del tiempo.¹¹⁰ La estructura básica es el objeto principal de la justicia política y social, ya que, principalmente, tales instituciones influirán durante toda la vida de un individuo. Esta estructura básica deberá estar regida por unos principios de justicia que ayudarán a que existan unos mínimos inamovibles que garanticen los medios suficientes para que cada ciudadano pueda llevar a cabo la concepción del bien que tenga establecida para su vida.

Otro de los elementos importantes que presenta Rawls como parte de su teoría es el de las facultades morales como base de la igualdad política. Estas facultades son: una concepción del bien y un sentido de la justicia. El individuo que quiere sumarse al proyecto

¹¹⁰ Cfr., John Rawls, *La justicia como equidad: una reformulación*, Paidós, Barcelona, 2000, pp. 31-33.

de la construcción de una sociedad equitativamente justa debe tener presente la importancia de la cooperación. La principal importancia de cooperar estriba en el hecho de no querer imponer una visión concreta del mundo que controle a los demás, ni creerse con mayores privilegios y prerrogativas por tener cierta cantidad de dinero o cualquier tipo de dotación innata que sobresalga sobre los demás. Antes bien, cooperar es trabajar, mancomunadamente, en la implementación de un proyecto de justicia social y política en el que todos salgan beneficiados.

Por lo tanto, poseer un sentido de la justicia se refiere a la capacidad de entender, aplicar y obrar según los principios de la justicia política y social. La concepción del bien se refiere, por su parte, a todos los fines y objetivos últimos que considere, todo individuo, para su plena realización.¹¹¹ El sentido de justicia, de igual manera, debe estar basado en la convicción de todo ciudadano de gozar de una libertad e igualdad que le concede el hecho de pertenecer a un sistema equitativo de cooperación durante toda su vida.

Hemos visto hasta aquí, someramente, el concepto de justicia como imparcialidad que presenta Rawls a lo largo de su teoría y, además, algunos de los principales elementos que la componen. Veamos ahora una de las polémicas propuestas de Rawls que, al mismo tiempo, es una respuesta a sus desacuerdos que con el utilitarismo; hablamos de los “bienes sociales primarios”. Analicemos la argumentación rawlsiana y de qué manera representa una distancia frente a la posición utilitarista clásica.

II. Bienes sociales primarios: una afrenta al utilitarismo

En una sociedad política bien ordenada, los ciudadanos presentarán necesidades que las instituciones políticas y sociales deben reconocer públicamente como tales, lo que, en consecuencia, resultará en beneficio de todos.¹¹² Además, como parte de estas necesidades se encuentran muchas concepciones comprensivas del bien; sin embargo, el gobierno, como en el caso del utilitarismo, no podría proceder de manera tal que lleve al máximo la satisfacción de las preferencias racionales de los ciudadanos, ni de sus necesidades.¹¹³ Ni tampoco podría ayudar a promover determinada concepción del bien sobre otras, como lo

¹¹¹ *Cfr.*, *Ibidem*, p. 43.

¹¹² *Cfr.*, John Rawls, *Liberalismo político*, Fondo de Cultura Económica, México, 1995, p. 176.

¹¹³ *Cfr.*, *Idem*.

habíamos señalado antes cuando hablábamos de la distinción entre una comunidad –o la imposición de una determinada doctrina comprensiva- y un pluralismo razonable. Como lo afirma Rawls, “tratar de realizar cualquiera de estos valores religiosos o filosóficos mediante la instauración de instituciones básicas daría a la sociedad política cierto carácter sectario”.¹¹⁴

Por consiguiente, en la idea de liberalismo político que intenta impulsar Rawls será importante buscar una solución al enfrentamiento de tantas convicciones diferentes y, muchas veces, distanciadas que existen entre los ciudadanos. Es por esto que la concepción política de la justicia ve la necesidad de proponer una lista de bienes primarios como solución práctica a este inconveniente. Estos bienes son los que, se supone, todos los ciudadanos necesitarán para garantizar un eficaz desarrollo de sus planes de vida, no importando las convicciones que los inspiren existencialmente.

Los bienes primarios son, por ende, “cosas que necesitan los ciudadanos en tanto que personas libres e iguales, y las exigencias acerca de estos bienes se consideran exigencias apropiadas”.¹¹⁵ A la lista de bienes primarios, Rawls no le cierra la posibilidad de agregar otros bienes, no obstante, afirma que los que ha propuesto son los que convergerían ante cualquier tipo de concepción del bien de un ciudadano convencido de ser un individuo libre e igual. La lista de bienes primarios que propone Rawls es la siguiente; a saber:

- a) Los derechos y libertades básicas;
- b) la libertad de desplazamiento y la libre elección de ocupación en un marco de diversas oportunidades;
- c) los poderes y las prerrogativas de los puestos y cargos de responsabilidad en las instituciones políticas y económicas de la estructura básica;
- d) ingresos y riquezas, y, por último,
- e) las bases sociales de respeto a sí mismo.¹¹⁶

Estos bienes primarios deben estar sustentados sobre un consenso traslapado que se deriva, directamente, del pluralismo razonable. Este consenso traslapado o consenso entrecruzado razonable consiste en el apoyo, por parte de cada ciudadano con sus diferentes concepciones del bien, de una misma concepción política de justicia para, de esta forma, sentar la base más

¹¹⁴ *Idem.*

¹¹⁵ *Ibidem*, p. 177.

¹¹⁶ *Idem.*

razonable de unidad política y social disponible para los ciudadanos de una sociedad democrática.¹¹⁷

La justicia como imparcialidad, a diferencia de la doctrina utilitarista, “rechaza la idea de comparar y llevar al máximo el bienestar general en materia de justicia política”.¹¹⁸ Las comparaciones interpersonales que hace Rawls están referidas a las necesidades que atañen, propiamente, a los ciudadanos. En este sentido, la justicia como imparcialidad se abstiene de promover un estilo de vida en concreto o un concepto de felicidad específico, y tampoco juzga los planes de vida de cada ciudadano; antes bien, cree que los bienes primarios especifican cuáles son las necesidades de los ciudadanos cuando se suscitan cuestiones acerca de la justicia,¹¹⁹ teniendo en cuenta que “la especificación de estas necesidades es una construcción que se hace desde el interior de una concepción política, y no desde el interior de una doctrina comprensiva”.¹²⁰

Como habíamos señalado antes, nuestro objetivo en este primer capítulo del trabajo era presentar los aportes de algunos de los principales autores utilitaristas clásicos para, de esta manera, vislumbrar la manera en que Rawls desarrolla su argumentación crítica frente a los lineamientos fundamentales de esta doctrina. Además, hemos analizado la propuesta de una justicia como imparcialidad, que da prioridad a la justicia sobre el bien, concibiéndola como la principal virtud de las instituciones sociales y políticas de la sociedad. Esto ameritaría, necesariamente, tomar distancia de la manera en que se venía considerando lo justo desde las convicciones utilitaristas.

En la siguiente parte abordaremos las respuestas de los utilitaristas contemporáneos a las críticas que hizo Rawls a la doctrina utilitarista clásica. En efecto, nos centraremos en la argumentación que hace John Harsanyi y la disputa que establece con la posición rawlsiana. Además, analizaremos algunas críticas que se le hacen tanto a Harsanyi como a Rawls. De esta forma, tendremos más elementos para analizar la viabilidad de la constitución de una sociedad justa a partir un modelo de posición original utilitarista.

¹¹⁷ *Cfr.*, *La justicia como equidad*, *Op. cit.*, p. 59.

¹¹⁸ John Rawls, *Liberalismo político*, *Op. cit.*, p. 183.

¹¹⁹ *Cfr.*, *Ibidem*, pp. 183-184.

¹²⁰ *Ibidem*, p. 184.

Capítulo 2. La posición original utilitarista desde el planteamiento ético de John Harsanyi

En este capítulo expondremos los principales argumentos del utilitarismo contemporáneo representado en la reflexión ética que hace John Harsanyi. Para ello, abordaremos su texto “Morality and the Theory of Rational Behaviour”,¹²¹ buscando resaltar las principales premisas de su principio ético que son imprescindibles dentro de las discusiones alrededor del utilitarismo y la justicia actualmente. Además, presentaremos la manera en que David Gauthier, en su libro *La moral por acuerdo*,¹²² problematiza las principales proposiciones de Harsanyi, así como las de John Rawls. Esto nos servirá para tener un punto de vista más amplio al momento de respondernos la pregunta acerca de la posibilidad de una sociedad justa bajo un modelo de posición original utilitarista.¹²³

1. John Harsanyi: el utilitarismo de la preferencia

El filósofo húngaro John Charles Harsanyi, nacido el 29 de mayo de 1920 en Budapest, es uno de los representantes más importantes de la doctrina utilitarista contemporánea. En 1994 le concedieron el Premio Nobel de Economía por sus trabajos sobre los análisis de los equilibrios en la teoría de juegos. Además de basar su propuesta en la teoría de juegos – a partir de la teoría moderna bayesiana de la conducta racional bajo riesgo e incertidumbre-¹²⁴ Harsanyi se guía, principalmente, por tres tradiciones morales en la elaboración de su teoría utilitarista. La primera tradición moral es la de Adam Smith acerca del punto de vista moral desde un observador imparcial y empático (*sympathetic*).¹²⁵ Otro de los planteamientos morales que guiarán la reflexión de Harsanyi es la argumentación kantiana a partir de la premisa de que “las reglas morales deben ser distinguidas de otras reglas conductuales a partir

¹²¹ John Harsanyi, “Morality and the Theory of Rational Behaviour”, *Social Research* 44, no. 4 (1977): 623-56. Accessed March 15, 2020. www.jstor.org/stable/40971169.

¹²² David Gauthier, *La moral por acuerdo*, Gedisa, Barcelona, 1986.

¹²³ En este capítulo no confrontaremos el utilitarismo contemporáneo con la propuesta de Rawls directamente. Buscaremos, más bien, identificar las principales premisas del utilitarismo contemporáneo desde la propuesta de Harsanyi. En el último capítulo es donde sacaremos nuestras conclusiones respecto del debate entre Harsanyi y Rawls.

¹²⁴ John Harsanyi, “Morality and the Theory of Rational Behaviour”, *Social Research* 44, no. 4 (1977): 623-56. Accessed March 15, 2020. www.jstor.org/stable/40971169, p. 623. Todas las traducciones son mías.

¹²⁵ *Ibidem*, p. 623.

de cierto criterio formal y, particularmente, desde el criterio de universalidad (el cual debe ser descrito, también, como un criterio de reciprocidad)”.¹²⁶

Pero la principal tradición moral que inspiró al economista húngaro fue el utilitarismo clásico de Bentham, John Stuart Mill, Sidgwick y Edgeworth. Estos autores, tal como explicamos en el capítulo anterior, tuvieron como base de su propuesta utilitarista la maximización de la utilidad social en tanto criterio de moralidad.¹²⁷ Harsanyi es consciente de que existen detalles particulares en las propuestas de los autores del utilitarismo clásico que no sería razonable aceptar en la contemporaneidad; no obstante, resalta el aporte significativo que estos filósofos hicieron tanto desde lo político como desde lo ético. En lo ético, el hecho de que propusieran “someter todas las reglas morales aceptadas a pruebas de racionalidad y utilidad social”.¹²⁸ Para el economista húngaro, en efecto, la herencia más importante que nos deja el utilitarismo clásico para la reflexión actual –la cual será el punto de partida de su propuesta utilitarista- es que “las cuestiones morales deben ser decididas por pruebas racionales y que la conducta moral, en sí misma, es una forma especial de conducta racional”.¹²⁹

Uno de los temas de discusión más relevantes en la reflexión tanto de la economía como de las ciencias políticas y la filosofía moral ha sido el de la teoría de la elección o la acción racional (*rational behaviour*). Para Harsanyi, el concepto de “elección racional” es importante en disciplinas normativas de teoría de la decisión, de teoría de juegos y, especialmente, en la ética. El autor define “elección racional” como “una acción que, consistentemente, persigue determinados objetivos, y lo hace con relación a un conjunto de preferencias o prioridades muy bien definidas”.¹³⁰

Ahora, no podemos negar que la conducta humana, al trabajar en la consecución de un determinado fin, tendrá siempre en el horizonte la incertidumbre de un actuar completamente racional. ¿Cómo sabremos que hemos actuado de tal forma? Harsanyi

¹²⁶ *Ibidem*, p. 624.

¹²⁷ *Idem*. Harsanyi anota que lo que los autores del utilitarismo clásico denominaron “utilidad social” equivale a lo que los estudios modernos de economía social llamarán “función del bienestar social” (*social welfare function*); no obstante, en muchos casos, este término es usado en un sentido menos específico, sin ninguna connotación utilitarista. Más adelante explicaremos lo que el autor define como la “función de bienestar social” en el contexto de su teoría.

¹²⁸ *Ibidem*, p. 625.

¹²⁹ *Idem*.

¹³⁰ *Ibidem*, p. 628.

responde a esta pregunta afirmando que, no obstante lo difícil de dirigir la conducta bajo una racionalidad ideal, existen formas de proceder de los individuos –tales como la vida económica, aspectos de la vida política, entre otras áreas de la interacción humana- que “muestran, suficientemente, altos grados de racionalidad como para dar una sorprendente cantidad de poder tanto explicativo como predictivo a varios modelos analíticos que ayudarían a postular una racionalidad completa (*full rationality*)”.¹³¹ Por ende, para el Nobel de economía, la tarea de las disciplinas normativas de teoría de la decisión, teoría de juegos y la ética consiste en “ayudar a las personas a actuar más racionalmente y darles un mejor entendimiento acerca de lo que, realmente, es la racionalidad”.¹³²

Harsanyi arguye que tanto la teoría de la decisión, la teoría de juegos como la ética son tres ramas de una misma disciplina: la teoría general de la elección racional. El economista define la “teoría de la decisión” como una teoría de la elección racional individual bajo riesgo (donde todas las probabilidades son objetivamente conocidas) e incertidumbre (donde algunas o todas las probabilidades son desconocidas, y podrían ser, incluso, vagamente definidas como probabilidades objetivas).¹³³ La teoría de juegos la define como “una teoría de interacción racional entre dos o más individuos, cada uno de los cuales persigue racionalmente sus propios objetivos contra otros individuos quienes, a su vez, persiguen racionalmente sus propios fines. Cualquier fin individual podría ser egoísta o desinteresado según lo determine su propia función individual”.¹³⁴ La ética es definida como “una teoría de elección racional al servicio de los intereses comunes de la sociedad en su conjunto”.¹³⁵

Hasta aquí hemos presentado, someramente, las principales bases del utilitarismo de Harsanyi. Ahora nos adentraremos en su teoría de la moralidad, que se fundamenta en un modelo de equiprobabilidad para los juicios de valor morales. Como economista que era, se valió del lenguaje matemático para explicar sus principales ideas; sin embargo, favoreciendo nuestro principal interés en esta investigación, no nos detendremos en los formalismos

¹³¹ *Idem.*

¹³² *Idem.*

¹³³ *Cfr., Ibidem*, p. 629.

¹³⁴ *Idem.*

¹³⁵ *Idem.*

matemáticos propios de su teoría. Antes bien, busquemos rastrear los principales argumentos de su propuesta ética para, luego, confrontarlos con la posición rawlsiana.

I. Modelo de equiprobabilidad

Para Harsanyi un “juicio de valor moral” es un juicio de preferencia. Se trata de preferencias que no se basan en criterios egoístas. Por ejemplo, es distinto un individuo que prefiere una sociedad X a una sociedad Y solo por intereses personales a un individuo que prefiere la sociedad Y sobre la sociedad X asumiendo que, en cualquier sociedad que haya elegido, tendrá la misma posibilidad de ocupar cualquiera de las posiciones sociales disponibles. Su elección sería independiente de cualquier tipo de consideraciones egoístas moralmente irrelevantes;¹³⁶ este sería, entonces, un juicio de valor genuinamente moral (*genuine moral value judgment*). Al hecho de asumir que se tendrá la misma probabilidad de ocupar cualquier posición social en una determinada elección, lo denomina Harsanyi “postulado de equiprobabilidad” (*equiprobability postulate*). Mientras que a todo el proceso de decisión basado en tal postulado se le denomina “modelo de equiprobabilidad de los juicios de valor morales” (*equiprobability modelo of moral value judgments*).¹³⁷

Supongamos, entonces, que en la sociedad existen n individuos (individuo 1,2,3..., n). Estos individuos podrían ocupar desde el puesto más bajo, pasando por los intermedios, hasta llegar al más alto en un sistema socialmente dado. Cada individuo disfrutará de cierto nivel de utilidad bajo ese sistema social elegido. Si un individuo i desea hacer el juicio de valor moral respecto de los méritos de dos sociedades determinadas, desde el postulado de equiprobabilidad, tal individuo asumirá que tiene la misma probabilidad $1/n$ de ocupar cualquier posición social y, por ende, le corresponderá cualquier nivel de utilidad.¹³⁸

Por consiguiente –siguiendo la teoría de decisión bayesiana- “un individuo racional siempre elegirá un determinado sistema social que maximice su utilidad esperada [...] representando el promedio aritmético del nivel de utilidad de todos los individuos de la sociedad”.¹³⁹ En otras palabras, afirma Harsanyi, “un individuo racional siempre usaría su

¹³⁶ Cfr., *Ibidem*, p. 632.

¹³⁷ Cfr., *Idem*.

¹³⁸ Cfr., *Idem*.

¹³⁹ *Ibidem*, p. 633.

utilidad promedio como su ‘función de bienestar social’, o sería un utilitarista que define la utilidad social como el promedio de la utilidad de todos los individuos (más que como su suma, como lo han hecho muchos utilitaristas)”.¹⁴⁰

En “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”,¹⁴¹ Harsanyi sostiene que “el concepto ingenuo de bienestar social, como una suma de utilidades cardinales intuitivamente medibles y comparables, ha sido encontrado incapaz de resistir el criticismo metodológico de la escuela de Pareto”;¹⁴² por lo que, basado en las ideas del profesor Abram Bergson,¹⁴³ el economista húngaro define así la “función de bienestar social”:

una función arbitraria de variables económicas (y otras sociales), de una forma libremente escogida de acuerdo con los juicios de valor éticos o políticos de una persona [...] Debido a la prevalencia de juicios de valor individualistas en nuestra sociedad, ha sido generalmente acordado que una función de bienestar social debería ser una función que incremente las utilidades de los individuos.¹⁴⁴

Harsanyi piensa en la existencia de postulados éticos como marco de una función de bienestar social. Esta idea la toma del profesor Fleming,¹⁴⁵ quien cree que los sistemas éticos “proveen una forma de medir el bienestar a través de un proceso de evaluación ética. Todo razonamiento que presuponga un sistema ético de este tipo “podría legítimamente hablar de

¹⁴⁰ *Idem.*

¹⁴¹ John Harsanyi, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *The Journal of Political Economy*, *Vo. 63, No. 4* (Aug. 1955), pp. 309-321, University of Chicago Press, Accessed March 12, 2020. <http://www.jstor.org/stable/1827/128>

¹⁴² *Ibidem*, p. 309. Todas las traducciones son mías.

¹⁴³ Abram Bergson (1914-2003) fue un economista estadounidense, “profesor de la universidad de Columbia, Texas y la universidad de Harvard. Su principal área de investigación fue la economía del bienestar (...) Definió y analizó el concepto de una función de bienestar social individualista. Esta función permite las condiciones necesarias para que un óptimo económico se calcule sin la hipótesis de utilidad cardinal. Por otra parte, la función se utiliza como una forma de priorizar y establecer un orden de preferencias entre las diferentes situaciones que cumplan los requisitos de Pareto para un óptimo social. Bergson postuló la necesidad de aceptar explícitamente juicios de valor como una manera de desarrollar funciones de indiferencia”. (*Cfr.*, <https://policonomics.com/es/abram-bergson/> Tomado el 09 de marzo de 2020).

¹⁴⁴ John Harsanyi, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *Op. cit.*, p. 309.

¹⁴⁵ John Marcus Fleming (1911-1976) fue un economista británico. Su principal propuesta ética consistió en que la función del bienestar social debe ser adicionalmente separable, por encima de los individuos, de los períodos de tiempo y de las cambiantes circunstancias del mundo. (Peter J. Hammond, *Harsanyi’s Utilitarian Theorem: A Simpler Proof and Some Ethical Connotations*, Department of Economics, European University Institute, Italy, Tomado de <https://web.stanford.edu/~hammond/HarsanyiFest.pdf>, el 23 de febrero de 2020).

bienestar en todo el sentido de una magnitud cardinal”.¹⁴⁶ De esta manera, a partir del análisis de uno de los postulados de Fleming,¹⁴⁷ en el cual se requiere que la distribución de la utilidad entre dos individuos deba siempre ser juzgada independientemente de la manera en que la utilidad y el ingreso están distribuidos entre otros miembros de la sociedad,¹⁴⁸ Harsanyi arguye que la propuesta de Fleming está basada en expectativas seguras sin dejar lugar a la incertidumbre. Por lo tanto, pensando en las preferencias definidas por elecciones entre perspectivas inciertas, el Nobel de economía propone una elección desde un punto de vista moral especificado a partir del punto de vista de un observador solidario y, a la vez, imparcial.¹⁴⁹ Esta elección se realizará bajo condiciones de riesgo e incertidumbre.

Harsanyi concibe la figura del “observador imparcial y empático” (*impartially sympathetic observer*), como una moderna reafirmación de la teoría de Adam Smith. En este caso, el requerimiento de imparcialidad de Smith corresponde al postulado de equiprobabilidad; mientras que su requerimiento de empatía corresponde a la asunción de Harsanyi de que el observador hará su elección en términos de una comparación interpersonal de utilidad basada en la empatía con los distintos miembros de la sociedad.¹⁵⁰

Harsanyi se basa en la teoría de juegos para plantear el punto de vista del observador imparcial; para ello, se sirve de la función de utilidad¹⁵¹ de von Neumann- Morgenstern, la cual consiste en proyectar, a través de axiomas básicos, la manera racional en que un

¹⁴⁶ Marcus Fleming, “A cardinal Concept of Welfare”, *The Quarterly Journal of Economics*, Vol. 66, No. 3 (Aug., 1952), pp. 366-384. Oxford University Press, Tomado de <https://jstor.org/stable/1885309>, p. 19. La traducción es mía.

¹⁴⁷ Este postulado, denominado “evaluación independiente de la distribución de la utilidad entre cada par de individuos”, reza que existen tres individuos, donde el individuo i es indiferente entre las situaciones X y $X1$ y también entre las situaciones Y y $Y1$; pero prefiere las situaciones X y $X1$ a las situaciones Y y $Y1$. No obstante, hay otro individuo j , que está en el mismo escenario que el individuo i ; no obstante, a diferencia de i , el individuo j prefiere Y y $Y1$ a X y $X1$. De igual manera, otros individuos son indiferentes entre X y Y , así como entre $X1$ y $Y1$. La preferencia social, en este caso, debe ir en la línea de una preferencia específica; es decir, si desde un punto de vista social se elige a X por encima tanto de los individuos, los períodos de tiempo y las cambiantes circunstancias del mundo, entonces se debería elegir a $X1$ también; lo mismo ocurriría en el caso de que se prefiriera a Y . (Cfr., *Ibidem*, p. 311).

¹⁴⁸ Cfr., *Ibidem*, p. 312.

¹⁴⁹ Cfr., *Ibidem*, pp. 312-313.

¹⁵⁰ Cfr., John Harsanyi, “Morality and the Theory of Rational Behaviour”, *Op. cit.*, p. 633.

¹⁵¹ Recordemos que una “función de utilidad es una ecuación matemática en la que se representa la ‘satisfacción’ o ‘utilidad’ que obtiene un consumidor cuando disfruta de una determinada cantidad de bienes o servicios” (Cfr., <https://economipedia.com/definiciones/funcion-de-utilidad.html>, consultado el 02 de marzo de 2020).

individuo elige desde la maximización de la utilidad esperada.¹⁵² Pensemos, por ejemplo, en diferentes sociedades o en la misma sociedad bajo diferentes políticas o reglas públicas, y denotemos estas diferentes reglas como A, B, C, D... etc. Cualquier posición objetiva de un individuo *i* respecto de cualquiera de estas políticas públicas se podría considerar como un vector que ayudaría a visualizar lo económico, lo social, lo biológico, entre otras variables.¹⁵³ Desde la “función del bienestar social”, para que el individuo *i* sea imparcial –desde un punto de vista moral y teniendo en cuenta la situación social desde la política pública A– debe partir de un escenario de total incertidumbre en cuanto a su posición determinada (*A_i*) frente a esta política, buscando cuál podría ser su lugar en A. Lo que propone Harsanyi es que el individuo *i*, como observador imparcial y empático, asigne la misma probabilidad ante la posibilidad de que a él le pueda tocar cualquier posición determinada en la sociedad.¹⁵⁴

La elección del mejor escenario, moralmente hablando, dependerá de optar por la maximización de la utilidad esperada. Podemos notar, nuevamente, que la propuesta del economista húngaro es similar a la doctrina clásica utilitarista. La novedad de Harsanyi radica en el hecho de que hace su planteamiento desde las variables de una función de utilidad como la de von Neumann Morgenstern, con lo que busca que las funciones de bienestar social no sean lineales en relación con los bienestar individuales; es decir, plantea concebir al bienestar social como un bienestar medio.¹⁵⁵ En este sentido, Harsanyi afirma que:

no podemos permitir la distinción entre una función individual de bienestar social y su función de utilidad si queremos sostener el principio de que una función de bienestar social debería estar basada no solo en la función de utilidad de un individuo en particular sino, más bien, en las funciones de utilidad -preferencias subjetivas- de todos los individuos, representando un tipo de ‘compromiso justo’ entre ellos.¹⁵⁶

¹⁵² El término “utilidad esperada” parte de la necesidad de “resolver el problema de definir la conducta aparentemente no racional de las personas frente a muchas situaciones” (Alexis Arechiles Moubarak, *Teoría de la utilidad esperada: una aproximación realista*, Universidad de Chile, Santiago, 2008, pp. 11-12); de manera que, al momento de hacer determinadas elecciones, los individuos contribuyan a la maximización de la utilidad esperada. La utilidad esperada vendría a tomar el lugar de lo que se conocía en economía como “valor esperado”. El término “maximización de la utilidad esperada” fue acuñado por los economistas John von Neumann y Oskar Morgenstern, precursores de la teoría de juegos.

¹⁵³ Cfr., Don Herzog, “Against utilitarianism”, *Justification in Political Theory*, Cornell University Press, 1985, Tomado de <https://www.jstor.org/stable/10.7591/j.ctt207g6t3.7>, p. 144. La traducción es mía.

¹⁵⁴ Cfr., *Ibidem*, p. 145.

¹⁵⁵ Cfr., “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *Op. cit.*, pp. 314-316.

¹⁵⁶ *Ibidem*, p. 315.

El economista húngaro habla de las “preferencias éticas” y las “preferencias subjetivas”. Con preferencias éticas se refiere al hecho de que, aunque la función de bienestar social y la función de utilidad, en cierto sentido, reflejen las preferencias de un determinado individuo, la función de bienestar social individual expresa lo que el individuo preferiría sobre la base de consideraciones sociales impersonales. Con preferencias subjetivas se refiere al hecho de que la función de utilidad de este individuo debe expresar qué es lo que realmente prefiere, tanto si lo hace basado en sus intereses personales como en cualquier otra motivación.¹⁵⁷ Además, agrega Harsanyi, sólo las preferencias subjetivas del individuo expresarán sus preferencias en todo el sentido de la palabra, tal como ellas realmente son, dejando al descubierto una actitud egoísta en el caso de un individuo egoísta, o una actitud altruista en el caso de uno altruista.¹⁵⁸ Las preferencias éticas de un individuo, entonces, expresarán lo que realmente llamaríamos “sus preferencias”, las cuales muestran lo que prefiere solo en los momentos en que se ve obligado a tener una actitud imparcial e impersonal sobre él mismo.¹⁵⁹

Las preferencias del individuo satisfacen el requerimiento de impersonalidad si estas indican qué situación social elegiría si no supiera cuál sería su posición social en una nueva situación elegida sino, más bien, siendo consciente de que tendrá la misma oportunidad de obtener cualquiera de las posiciones sociales existentes en esa nueva situación elegida.¹⁶⁰ Por lo tanto, agrega Harsanyi, es irrelevante que el individuo desconozca la manera en que su elección afectaría sus intereses personales o, simplemente, que ignore el conocimiento que tenga sobre estos intereses al momento de elegir. Hay que acotar que esta elección impersonal de la que habla el economista húngaro se refiere a una elección entre expectativas inciertas.¹⁶¹

Ahora, como habíamos mencionado antes, estas preferencias individuales, en cuanto impersonales y, a la vez, racionales, estarían definiendo “una función de bienestar social cardinal igual al promedio aritmético de utilidades de todos los individuos en la sociedad”.¹⁶² Es decir, si un individuo tiene un determinado criterio objetivo para comparar sus utilidades frente a las de los demás individuos, entonces su función de bienestar social representará un promedio indeterminado (*un-weighted mean*) de esas utilidades; mientras que sin tal criterio

¹⁵⁷ Cfr., *Ibidem*, p. 315.

¹⁵⁸ Cfr., *Idem*.

¹⁵⁹ Cfr., *Idem*.

¹⁶⁰ Cfr., *Ibidem*, p. 316.

¹⁶¹ Cfr., *Idem*.

¹⁶² *Idem*.

objetivo existirá, en efecto, un promedio ponderado (*weighted mean*), con una importancia arbitraria dependiendo solo de sus juicios de valor personales.¹⁶³ Este argumento ayuda a entender un poco más la postura ética de Harsanyi, desde la que los juicios éticos de una persona son la expresión no tanto de su actitud subjetiva, sino, de ciertas actitudes impersonales e imparciales.

II. Comparaciones interpersonales de utilidad y el postulado de similitudes en Harsanyi

Otro de los puntos que aborda Harsanyi está relacionado con las comparaciones interpersonales de utilidad. Es un tema que hace parte de nuestra vida diaria; cuando, por ejemplo, queremos regalar un libro que ya no leeremos y sopesamos a quién, de nuestro círculo cercano, le sería más grato disfrutar de tal libro. Puede que sea un libro de *thriller* juvenil que nuestros abuelos no disfrutarían mientras que sí lo haría un joven sobrino nuestro. Así existen otras muchas situaciones donde nos vemos abocados a hacer estas comparaciones de utilidad entre varios individuos.

Además, las comparaciones interpersonales de utilidad son un tema al cual todo autor utilitarista debería hacerle frente, ya que es esencial en cualquier propuesta ética utilitarista contemporánea y, asimismo, será relevante al momento de hablar sobre los términos en la construcción de justicia. El principal problema que se ha debatido respecto de estas comparaciones radica en la base que debe sostenerlas en el marco de una teoría ética.¹⁶⁴ Veamos la manera en que Harsanyi aborda esta cuestión.

Las comparaciones interpersonales se presentan tanto en forma de comparaciones de la satisfacción total de cada persona como de la comparación del aumento o disminución de la satisfacción de cada individuo. Para llevar a cabo las comparaciones interpersonales, según Harsanyi, existen dos indicadores de la utilidad que los individuos relacionan con diferentes situaciones. Por un lado, están las preferencias representadas en las elecciones de cada persona y, por otro lado, las expresiones de satisfacción o insatisfacción en cada situación.¹⁶⁵ Según el economista, no se presentaría ningún problema si un individuo usa estos indicadores

¹⁶³ Cfr., *Idem*.

¹⁶⁴ Cfr., *Ibidem*, p. 317.

¹⁶⁵ Cfr., *Idem*.

para realizar una comparación de las utilidades atribuyéndolas a diferentes situaciones. El problema se presentaría al momento de usar los indicadores de la utilidad para comparar la utilidad que diferentes personas atribuyen a cada situación. Harsanyi explica este problema desde dos ángulos: uno metafísico y otro psicológico.¹⁶⁶

El problema metafísico se presentaría “si tratáramos de comparar las utilidades disfrutadas por diferentes personas con idénticas preferencias y con idénticas reacciones a cualquier situación”.¹⁶⁷ Harsanyi denomina “principio de diferenciación injustificada” (*principle of unwarranted differentiation*) al hecho de que en el caso en que dos individuos muestren conductas similares en todos sus aspectos relevantes abiertos a la observación, cualquier conclusión a la que se llegue con respecto a alguna diferencia –no observable– sería considerada una hipótesis contraria a la lógica inductiva del método científico.¹⁶⁸ En ese caso, entraríamos a discutir aspectos como el estado mental de los individuos, sus distintas susceptibilidades a la satisfacción, entre otros aspectos que harían de las comparaciones interpersonales algo bastante confuso.

Sidgwick afirma que en la comparación interpersonal de utilidad el individuo estima y compara sus propios niveles de felicidad basándose en su propia introspección y memoria, sabiendo que el placer y el dolor son signos de experiencias agradables o desagradables, respectivamente.¹⁶⁹ Sin embargo, este planteamiento es insuficiente al momento de abordarlo desde el análisis que hace Harsanyi desde el ángulo metafísico.

La dificultad psicológica de las comparaciones interpersonales de utilidad está basada en la “cuestión acerca de cómo diferencias psicológicas entre las personas, en el amplio sentido de la palabra – hábitos de consumo, trasfondo cultural, estatus social, así como diferencias psicológicas heredadas o adquiridas- afectan la satisfacción que los individuos obtienen de cada situación”.¹⁷⁰ Por ejemplo, imaginemos que entre dos individuos, José y María, José prefiere leer libros por medios electrónicos que leerlos en formato físico; mientras que María prefiere lo segundo, esto es, tener la experiencia de leer un libro que no sea de formato digital. Esta distinción de preferencias se podría dar porque, José encuentra

¹⁶⁶ Cfr., *Idem*.

¹⁶⁷ *Idem*.

¹⁶⁸ Cfr., *Idem*.

¹⁶⁹ Cfr., *Lecciones sobre la historia de la filosofía política, Op. cit.*, p. 475.

¹⁷⁰ John Harsanyi, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *Op. cit.*, p. 318.

mayor utilidad en leer desde su celular o cualquier otra plataforma digital y, por lo tanto, no existen muchas utilidades para él leyendo en el modo tradicional. Ahora, al momento de comparar la utilidad entre las preferencias de María y José, ¿cómo sabremos el nivel de satisfacción que siente José al usar sus *gadgets* para leer libros, y el de María al sumergirse en su lectura con un libro físico en sus manos? Además, ¿hasta qué punto María y José sienten satisfacción o insatisfacción? Estas interrogantes muestran la insuficiencia de hacer comparaciones desde parámetros psicológicos: “en general, entre más grandes son las diferencias psicológicas, biológicas, culturales, y sociales entre dos individuos, mayor será el margen de error relacionado con las comparaciones entre sus utilidades”.¹⁷¹

Harsanyi denomina “postulado de similitudes” (*similarity postulate*) al hecho de que asumamos que diferentes personas tienen los mismos sentimientos psicológicos acerca de cualquier situación; sería imposible verificar tal supuesto desde una observación directa, así como tampoco podríamos conocer la naturaleza real de lo que están sintiendo tales individuos.¹⁷² Este postulado, por lo tanto, lleva a que el individuo no se sustraiga de los rasgos que lo hacen ser persona en sociedad con otros. Esto evitaría caer en solipsismos u otras experiencias donde los otros individuos no tendrían absolutamente nada en común con el individuo en cuestión. Antes bien, esto nos llevaría a darnos cuenta de que “vivimos realmente en un mundo poblado por millones de otros seres humanos, tan reales y tan conscientes como nosotros mismos”.¹⁷³

Entre el punto metafísico y psicológico, Harsanyi afirma que este último no ha sido totalmente rechazado. A pesar de la complejidad que conlleva, las variables psicológicas nos podrían aportar información relevante acerca de ciertas características de los individuos al momento de llevar a cabo comparaciones de utilidad: “las comparaciones interpersonales de utilidad no son juicios de valor basados en postulados éticos o políticos, sino que, más bien, consisten en proposiciones basadas en proposiciones factuales que se desprenden de principios de lógica inductiva”.¹⁷⁴

¹⁷¹ *Ibidem*, pp. 318-319.

¹⁷² *Cfr.*, “Morality and the Theory of Rational Behaviour”, *Op. cit.*, p. 640.

¹⁷³ *Ibidem*, p. 641.

¹⁷⁴ John Harsanyi, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *Op. cit.*, pp. 319-320.

No obstante, el Nobel de economía no desecha la posibilidad de que las comparaciones interpersonales estén, de alguna forma, basadas en juicios de valor éticos o políticos. Por ejemplo, el hecho de considerar que todos los individuos tienen la misma capacidad de satisfacciones parte del juicio de valor basado sobre la igualdad de trato que merecen todos los seres humanos; observamos, en este caso, que la afirmación sobre la susceptibilidad a la satisfacción de todos los individuos no depende de una creencia acerca de la igualdad de rasgos psicológicos entre las personas sino, más bien, de un derecho que se presenta como ideal para todos los individuos.¹⁷⁵

A partir del análisis que realiza desde las comparaciones interpersonales y de la propuesta alrededor de una ética utilitarista, una de las conclusiones de Harsanyi es que, ciertamente, aunque no tengamos el suficiente cuidado y pericia haciendo comparaciones interpersonales en el diario vivir, sabemos que necesitamos hacerlas en orden a tomar decisiones morales en cualquier situación específica de la vida. No podemos, por ejemplo, quedarnos toda la vida buscando cuál de nuestros familiares o amigos necesitaría, realmente, el libro que deseamos regalar; tendremos, necesariamente, que tomar una decisión.

Además, el hecho de que estemos dispuestos a hacer comparaciones interpersonales de utilidad desde un postulado de similitudes, quiere decir que somos conscientes de que vivimos con otros seres humanos, que son tan reales como nosotros y con los cuales compartimos una misma humanidad. De igual modo, esos otros individuos también tienen la misma capacidad de satisfacción o insatisfacción que nosotros poseemos, a pesar de las innegables diferencias individuales que, inevitablemente, siempre existirán.¹⁷⁶ El postulado de similitudes se presenta también como una forma de hacer frente a una lógica inductivista radical aplicada a estas comparaciones; de ahí que Harsanyi afirme que “la necesidad práctica de establecer políticas públicas requerirá que nos formulemos funciones de bienestar social—explícita o implícitamente— incluso si carecemos de una información factual necesaria para establecer una comparación interpersonal de utilidad sobre una base objetiva”.¹⁷⁷

¹⁷⁵ *Cfr.*, *Ibidem*, p. 320.

¹⁷⁶ *Cfr.*, “Morality and the Theory of Rational Behaviour”, *Op. cit.*, p. 641. Vemos, pues, que el problema de una base para las comparaciones interpersonales queda sin solución desde el edificio argumentativo de Harsanyi. En el próximo capítulo abordaremos la crítica de Rawls a partir de esta premisa.

¹⁷⁷ John Harsanyi, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *Op. cit.*, p. 320.

Hasta aquí hemos abordado los principales matices de la propuesta utilitarista de John Harsanyi. Hemos analizado, principalmente, cómo su propuesta ética se desmarca del utilitarismo clásico, un utilitarismo hedonista donde tanto la utilidad social como la función de utilidad individual son definidos en términos de placer y dolor; mientras que el utilitarismo que plantea el economista húngaro está construido desde las preferencias personales de las personas, valiéndose de la figura del observador imparcial y empático que garantiza la maximización de la utilidad esperada, asegurando, de este modo, la mayor utilidad a todos los miembros de la sociedad. Veamos ahora algunas de las críticas que se le hacen al utilitarismo contemporáneo desde la propuesta del filósofo David Gauthier.

2. Principales críticas al utilitarismo contemporáneo: Gauthier frente a Harsanyi y Rawls

Dentro de la amplia discusión que existe sobre las propuestas utilitaristas contemporáneas nos centraremos en algunas de las críticas más relevantes que se le hacen a Harsanyi, por tratarse de una de las voces más representativas del planteamiento ético utilitarista contemporáneo. Conocer tales críticas nos dará recursos para, en el siguiente capítulo, contrastar la propuesta de Harsanyi con la teoría rawlsiana desde el escenario de la posición original; buscando rastrear así alguna pista –si la hubiera- que permita hablar de una sociedad imparcialmente justa bajo un modelo de posición original utilitarista.

Para el propósito descrito tomaremos como autor central a David Gauthier (1932), quien plantea una teoría moral neocontractualista partiendo de la teoría de la elección racional, así como de fuertes críticas tanto al utilitarismo de Harsanyi como a la teoría de la justicia de Rawls. Iniciaremos exponiendo las principales críticas a Harsanyi para, en un segundo momento, reparar en la manera en que Gauthier se enfrenta a Rawls. Aunque no nos detendremos ampliamente en su teoría moral, sí presentaremos algunos de los principales aspectos de esta.

En su obra magna, *La moral por acuerdo*¹⁷⁸, Gauthier tratará de responder a la pregunta de si los deberes morales están realmente basados en la racionalidad. Mostrará que la razón desempeña un papel práctico que está relacionado con el interés individual y que

¹⁷⁸ Gauthier, David, *La moral por acuerdo*, Gedisa, Barcelona, 1986.

además lo trasciende, de tal modo que los principios morales¹⁷⁹ que prescriben deberes, pasando por alto las ventajas o beneficios, pueden también justificarse racionalmente. El autor sostendrá, asimismo, que la concepción tradicional de la moral es una limitación en la búsqueda del interés personal.¹⁸⁰

Con su teoría, Gauthier tratará de desarrollar la conexión profunda que existe entre la moral y la elección racional, lo que no logran, en su opinión, ni Rawls ni Harsanyi; éstos no consideran los principios morales como un subconjunto de principios racionales necesarios para la elección.¹⁸¹ Los principios que propone Rawls, por ejemplo, no están directamente relacionados con la toma de decisiones individuales.¹⁸² Por ende, Gauthier se propone “mostrar por qué un individuo, al razonar partiendo de premisas no morales, estaría dispuesto a aceptar en sus elecciones las restricciones de la moral”.¹⁸³ En otras palabras, su interés es “validar la concepción que entiende la moral como un conjunto de restricciones racionales e imparciales que limitan la búsqueda del interés personal y no defender ningún código moral particular”.¹⁸⁴

En el complejo mundo moral del individuo normalmente existe un lugar en que éste se apoya para considerar sus decisiones como buenas o malas moralmente. A este lugar le denomina Gauthier “punto arquimédico”, desde el cual la persona tiene la capacidad moral de dar forma a la sociedad.¹⁸⁵ Tal punto lo caracteriza Gauthier en referencia a lo que es la persona y no desde características o circunstancias exteriores. El ideal de un individuo ubicado en el punto arquimédico es éste:

un actor racional liberado, no de la individualidad sino del contenido de cualquier individualidad particular, un actor consciente de que es un individuo con facultades y preferencias, particulares en sí mismas y distintivas respecto de las de los demás, pero que no reconoce cuáles son esas facultades y preferencias.¹⁸⁶

¹⁷⁹ Para Gauthier “los principios morales” son “los principios racionales necesarios para hacer una elección o para tomar decisiones dentro de las acciones posibles, los cuales incluyen algo que restringe de manera imparcial la acción del actor que persigue su propio interés”. (*Ibidem*, p. 17).

¹⁸⁰ *Cfr.*, *Ibidem*, p. 16.

¹⁸¹ *Cfr.*, *Ibidem*, p. 19.

¹⁸² *Cfr.*, *Idem*.

¹⁸³ *Ibidem*, p. 20.

¹⁸⁴ *Ibidem*, p. 21.

¹⁸⁵ *Cfr.*, *Ibidem*, p. 308.

¹⁸⁶ *Idem*.

Es decir, Gauthier parte de la capacidad ineludible –de un actor ideal ubicado en el punto arquimédico– de interactuar con otros individuos y, por ende, de poseer éste un interés imparcial por tal interacción se verá conducido a pensar en la posibilidad de diferentes configuraciones de estructuras sociales. Del mismo modo, siendo el actor ideal un sujeto plenamente racional será, por lo tanto, imparcial y, en efecto, los principios que elija para la sociedad expresarán normas de justicia; además, siendo una persona plenamente informada, elegirá una estructura de sociedad justa.¹⁸⁷ En consecuencia, un actor situado en el punto arquimédico ayudaría a llegar a un acuerdo racional sobre ciertas estructuras sociales, instituciones y prácticas- como el mejor camino para determinar condiciones sociales de convivencia.¹⁸⁸

El actor ideal es considerado un maximizador de la utilidad que buscará obtener el mayor beneficio. Sus utilidades pueden ser independientes de las utilidades de los demás; en este sentido, el punto arquimédico es compatible con el supuesto de “mutuo desinterés”.¹⁸⁹ El velo de la ignorancia que atribuye Gauthier a su actor ideal radica en el hecho de que no puede identificarse con una persona en particular dentro de la sociedad; sin embargo, tal actor tiene un conocimiento completo de la sociedad y sus miembros. Ese actor posee tal conocimiento y tales capacidades en tanto está ubicando en un punto arquimédico; no sabrá, con antelación, qué clase de sociedad le tocará vivir antes de llevar a cabo su elección; más bien, sólo debe tener en cuenta “las formas en las cuales las preferencias, las facultades y las circunstancias pueden encajar dentro de las sociedades factibles”.¹⁹⁰

Cualquier individuo se podría identificar con el actor ideal, ya que una de sus funciones será maximizar la utilidad de la persona que le toque ser cuando haya elegido una estructura determinada de sociedad. Esta identificación se da, también, por la imparcialidad que subyace a tal proceder: “la imparcialidad de la elección se encuentra no en ausencia de la preocupación por aquellos que resulten afectados, sino en presencia de una misma preocupación, y esta última queda asegurada por el objetivo maximizador del actor ideal”.¹⁹¹

¹⁸⁷ *Cfr., Ibidem*, p. 309.

¹⁸⁸ *Idem*.

¹⁸⁹ *Cfr., Ibidem*, P. 311.

¹⁹⁰ *Ibidem*, p. 312.

¹⁹¹ *Idem*.

Un punto en el que concuerdan Gauthier y Harsanyi es en el planteamiento que hace este último de la imparcialidad: “los principios de Harsanyi son estrictamente hipotéticos, rigen la elección racional partiendo de una posición imparcial o suponiendo preferencias imparciales, con lo cual sólo son principios válidos para alguien que quiere elegir moral o imparcialmente”.¹⁹² No obstante, Harsanyi no pretende “que haya situaciones en las que un individuo deba elegir moralmente para elegir racionalmente”,¹⁹³ escenario al cual logra llegar Gauthier. Específicamente, en Harsanyi existe un modo racional de elegir moralmente; mientras que en Gauthier existe una exigencia racional para la elección moral.

La principal crítica que le hace Gauthier a Harsanyi se sigue del anterior argumento. El observador imparcial del economista húngaro deberá elegir los principios para una determinada estructura de sociedad que le garantice un promedio mayor de utilidad esperada. Como habíamos señalado antes, este promedio lo sustenta Harsanyi desde las comparaciones interpersonales de utilidad. Gauthier, en cambio, propone otra forma de hacer comparaciones interpersonales de utilidad, partiendo de la premisa de que “Harsanyi se equivoca al suponer que su procedimiento de elección es una simple extensión de la elección racional realizada en circunstancias de incertidumbre”.¹⁹⁴ Veamos cómo desarrolla esta crítica.

Retomando la teoría de Harsanyi, el actor ubicado en el punto arquimédico, al haber comparado las utilidades de todos los individuos, podrá establecer tanto la mayor utilidad como el mayor grado de satisfacción de la preferencia. Ahora, al conocer las utilidades esperadas de cada individuo en una determinada estructura social, lo que ayudará a determinar el promedio de la satisfacción de la preferencia en tal estructura será el promedio de las utilidades esperadas de los individuos.¹⁹⁵ Cuando hablamos de utilidad esperada, como habíamos anotado antes, estamos haciendo referencia a un cálculo acerca de hasta qué punto se podrían satisfacer las preferencias de los individuos. Entonces, la pretensión de Harsanyi es que se maximice la utilidad esperada promedio en un medio de incertidumbre y riesgo, y este punto es el que le critica Gauthier.

Para Gauthier, si el razonamiento de Harsanyi fuese acertado, podríamos afirmar que cualquier persona en la sociedad se identificaría con el actor ideal sabiendo que le

¹⁹² *Ibidem*, p. 20.

¹⁹³ *Idem*.

¹⁹⁴ *Ibidem*, p. 316.

¹⁹⁵ *Cfr.*, *Ibidem*, p. 317.

correspondería la maximización de la utilidad esperada promedio. No obstante, según Gauthier, “puede caracterizarse cada estructura social como una lotería en la cual todos los premios tienen las mismas probabilidades y cada premio resulta ser una persona con determinadas facultades y preferencias, que se encuentra en una particular situación social”.¹⁹⁶ Ahora, asignémosle a cada premio una utilidad; es decir, la utilidad de un individuo bajo cierta situación. Para Harsanyi, desde esta lógica, a cualquier persona x –que no conoce qué lugar le tocará en la sociedad ni, por lo tanto, su identidad– le corresponderá el promedio de las utilidades de todas las personas bajo sus situaciones específicas.

Por ejemplo, imaginemos que una sociedad está conformada por Pedro, Danilo y Mariela, los cuales viven bajo las situaciones, respectivamente, M, N y L. La utilidad esperada para cualquier individuo x será, entonces, el promedio de la suma de la utilidad de las situaciones de cada una de las tres personas: $U_x = 1/3 [U(\text{Pedro}) M + U(\text{Danilo}) N + U(\text{Mariela}) L]$, donde U es la utilidad. Gauthier afirma que aquí hay algo que falla desde el razonamiento de Harsanyi; a saber: el individuo x no recibirá, como lo insinúa Harsanyi, el premio $U_p(M)$ –la utilidad de Pedro en la situación M–, sino $U_x(P, M)$, es decir, la utilidad que obtendría x si Pedro estuviera en la situación M. Además, si comparamos interpersonalmente las utilidades nos daría que la $U_p(M) = U_x(P, M)$; es decir, la utilidad de Pedro en la situación M equivaldría a la utilidad del individuo x encontrándose Pedro ubicado en la situación M. Por ende, para Gauthier esta conclusión sería incorrecta si y solo si el grado de satisfacción de las preferencias que obtiene x estando P en M es el mismo que obtiene estando ella misma (x) en M.¹⁹⁷

En otras palabras, “el actor ideal, al maximizar la utilidad esperada promedio, no está maximizando la utilidad esperada, ni la propia ni la de ningún otro”.¹⁹⁸ Esto supondría pensar, según Gauthier, que un individuo no estaría eligiendo imparcialmente, sino que “ha de exhibir un completo desinterés por conocer su propia identidad, sus facultades y sus preferencias y que sólo ha de preocuparse por aquellas facultades que le permitan satisfacer sus preferencias”.¹⁹⁹ Es decir, la persona que se identifica con el actor ideal ubicado en el punto arquimédico es consciente de que la elección del actor ideal es la misma que hubiese

¹⁹⁶ *Ibidem*, p. 318.

¹⁹⁷ *Cfr.*, *Ibidem*, p. 318-319.

¹⁹⁸ *Ibidem*, p. 321.

¹⁹⁹ *Ibidem*, p. 322.

hecho encontrándose bajo un escenario ideal, porque “una elección es racional sólo para aquellos capaces de identificarse con ella y es imparcial sólo si todos pueden identificarse con ella”.²⁰⁰

Además, el utilitarismo promedio no presenta una base para la elección que hará el actor ideal, tal como le critica Rawls al utilitarismo. Gauthier agrega que el utilitarismo, de igual forma, no toma en serio la individualidad de las personas; es decir, del utilitarismo se podría desprender –además de que una persona deba sacrificar su utilidad por aumentar la de los demás– que se use a una persona como un medio para lograr la satisfacción de la preferencia general.²⁰¹ En palabras de Gauthier: “el utilitarismo viola la integridad del individuo como ser humano con sus propias facultades y preferencias características y, por lo tanto, con una utilidad distintiva que no es intercambiable con las utilidades de los demás que él trata de maximizar”.²⁰²

A pesar de este tipo de críticas señaladas que pueden desprenderse de propuestas como la de Harsanyi, también es cierto, para Gauthier, que si la elección que se hará desde el punto arquimédico debiera representarse en términos de maximización, la mejor opción sería el utilitarismo. Sin embargo, como hemos constatado, elegir buscando maximizar la utilidad esperada desconocería “un orden que responda positivamente a los órdenes de preferencia y que evalúe del mismo modo el orden de cada individuo”.²⁰³ Este es el problema que trata de resolver Gauthier y para ello establece una pregunta; a saber: “¿Cómo podemos representar la elección hecha desde el punto arquimédico, de modo tal que cada persona, consciente de sus propias facultades y preferencias, sea capaz de identificarla con la elección que haría si no fuese consciente de ellas?”.²⁰⁴

Una opción viable para responderse a tal pregunta la encuentra Gauthier en el planteamiento de John Rawls; especialmente, en la manera en que el escenario que éste propone para que los “representantes ideales” elijan tras el “velo de ignorancia” conlleva a que cualquier individuo pueda identificarse con las elecciones que hagan tales representantes

²⁰⁰ *Idem.*

²⁰¹ *Cfr., Idem.*

²⁰² *Idem.*

²⁰³ *Ibidem*, pp. 322-323.

²⁰⁴ *Ibidem*, p. 123.

desde la posición original. Gauthier analiza, específicamente, el “principio de diferencia”²⁰⁵ de Rawls, a partir de la teoría de los “bienes primarios”,²⁰⁶ los cuales responden a la pregunta que plantea Rawls acerca de las comparaciones interpersonales de bienestar que hace el utilitarismo.²⁰⁷

Para Gauthier, de la propuesta rawlsiana se sigue que “el individuo debe tener especial cuidado de evitar lo peor que pueda acontecer”,²⁰⁸ de igual forma, “los términos de la cooperación, en la medida en que beneficien a algunos más que a otros, deben hacerlo de un modo que concite el apoyo voluntario de aquellos menos favorecidos”.²⁰⁹ Además, en Rawls estaría la idea de que si el menos favorecido se identifica con la elección arquimédica, todos pueden hacerlo y, por lo tanto, “tal elección satisface las exigencias de igual racionalidad y de imparcialidad”.²¹⁰

La crítica de Gauthier a Rawls se ubica en el hecho de que este último considere los los haberes naturales como un producto social. Según la argumentación rawlsiana, “el principio de diferencia representa, en efecto, el acuerdo de considerar la distribución de talentos naturales, en ciertos aspectos, como un acervo común, y de participar en los beneficios de esta distribución, cualesquiera que sean”.²¹¹ Gauthier considera que “Rawls exagera dramáticamente la influencia que ejercen las particulares circunstancias históricas y sociales en la formación de los seres humanos”.²¹² Esto lo lleva a considerar que Rawls cae en un “instrumentalismo social parcial” al pensar que los factores que determinan la elección

²⁰⁵ Este principio establece que “las desigualdades sociales y económicas tienen que satisfacer dos condiciones: en primer lugar, tienen que estar vinculadas a cargos y posiciones abiertos a todos en condiciones de igualdad equitativa de oportunidades; y, en segundo lugar, las desigualdades deben redundar en un mayor beneficio de los miembros menos aventajados de la sociedad” (John Rawls, *La justicia como equidad: una reformulación*, Paidós, Barcelona, 2000, p. 73).

²⁰⁶ Para Rawls, los “bienes sociales primarios” son las condiciones sociales y los medios de uso universal necesarios para que los ciudadanos se desarrollen adecuadamente y ejerzan plenamente sus dos facultades morales y, de igual manera, promuevan sus concepciones específicas del bien. (Cfr., *Ibidem*, p. 90).

²⁰⁷ En el próximo capítulo profundizaremos en esta crítica que le hace Rawls al utilitarismo de John Harsanyi.

²⁰⁸ David Gauthier, *La moral por acuerdo*, *Op. cit.*, p. 325.

²⁰⁹ *Ibidem*, p. 325.

²¹⁰ *Idem*. Nozick, en *Anarquía, estado y utopía*, diría jocosamente –parafraseando a Rawls– que el individuo menos favorecido se dirigiría al más favorecido de la siguiente manera: “Miren, mejor dotados: ustedes ganan cooperando con nosotros. Si quieren nuestra cooperación, tendrán que aceptar condiciones razonables. Sugerimos estas condiciones: cooperaremos con ustedes si nosotros obtenemos tanto como sea posible. Esto es, las condiciones de la cooperación deben darnos esa máxima porción de manera que, si se tratara de darnos más, terminaríamos con menos”. (Robert Nozick, *Anarquía, estado y utopía*, Fondo de Cultura Económica, México, 1974, p. 194).

²¹¹ John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 104.

²¹² David Gauthier, *La moral por acuerdo*, *Op. cit.*, p. 331.

y la preferencia son arbitrarios; esto supondría, en efecto, un divorcio entre la moral y las condiciones de la acción racional.²¹³ Por ende, según el razonamiento de Gauthier, la postura de Rawls estaría en contra de la imparcialidad misma: “no se alcanza la imparcialidad considerando que las personas carecen de facultades e intereses particulares; en realidad, semejante enfoque viola la imparcialidad entre los individuos reales”.²¹⁴

Por consiguiente, para Gauthier, Rawls estaría cayendo en el mismo error en que caen los utilitaristas, esto es, “no tomar en serio la individualidad de las personas”.²¹⁵ La individualidad de las personas, para Gauthier, se basa en sus facultades y preferencias, y parece que Rawls les resta importancia en aras de la instrumentalización social que se desprende de su teoría. Esta crítica a Rawls le ayuda a Gauthier a seguir construyendo la manera en que la moral y la racionalidad podrían comulgar bajo la figura de un actor ideal ubicado en un punto arquimédico. Es así como, desde la crítica descrita, el autor canadiense considera que “el individuo real que debe identificarse con la elección ideal tiene sus propias características particulares y definitorias: facultades, talentos, actitudes, preferencias”.²¹⁶

Para efectos de nuestro objetivo en este trabajo, no seguiremos profundizando en el desarrollo de la confrontación de Gauthier tanto con Harsanyi como con Rawls. Nuestra principal intención al incursionar en la teoría del primero era alimentar la discusión en torno a estos últimos. De esta manera, tendremos más herramientas para adentrarnos en la forma en que Rawls concibe su propuesta desde el escenario de la posición original, así como para examinar la posibilidad de construir una sociedad imparcialmente justa desde una posición original utilitarista.

Capítulo 3. Configuración de la PO rawlsiana: imposibilidad de una sociedad imparcialmente justa desde una PO utilitarista

En los capítulos anteriores hemos analizado los principales postulados del utilitarismo desde su versión clásica hasta sus posturas contemporáneas más influyentes. Además, hemos

²¹³ *Cfr., Ibidem*, p. 333.

²¹⁴ *Idem*.

²¹⁵ *Idem*.

²¹⁶ *Ibidem*, p. 335.

presentado, someramente, algunas de las críticas más destacadas de Rawls al utilitarismo clásico. Este recorrido nos ha dado pistas para ir identificando los ejes fundamentales en la discusión utilitarista frente a la propuesta rawlsiana. En este capítulo nos proponemos profundizar en algunas de las críticas –apenas mencionadas en el primer capítulo- que hace Rawls al utilitarismo; especialmente, en los puntos cruciales de la discusión entre el utilitarismo de Harsanyi y la teoría de la justicia como imparcialidad. Para ello, nos ubicaremos en el escenario de la posición original (PO),²¹⁷ desde el planteamiento de ambos autores, buscando evaluar la posibilidad de una sociedad imparcialmente justa bajo el modelo de posición original utilitarista.

1. Posición Original rawlsiana

A partir de los principales parámetros de la justicia como imparcialidad explicados en el capítulo anterior, y luego de examinar el planteamiento del modelo de equiprobabilidad de Harsanyi representado en la figura del “observador imparcial”, nos adentraremos en el análisis de los cimientos del edificio rawlsiano, esto es, en la manera en que Rawls dispone las bases para sugerir un estilo de sociedad imparcialmente justa. La PO nos llevará a distinguir el modo en que Rawls elige una concepción de justicia que conduce a la construcción de un sistema justo de cooperación entre ciudadanos libres e iguales. Esto nos servirá de ayuda al momento de evaluar la posibilidad de un criterio utilitarista rector en una sociedad equitativamente justa.

Lo primero que hay que preguntarse es el por qué Rawls ve la necesidad de acudir a una PO; para responder a tal interrogante tenemos que volver a varios de los elementos cruciales de la teoría rawlsiana sobre los que habíamos hecho mención. Nos referimos a la cooperación social, específicamente, a los términos de ésta. Respecto de lo cual, Rawls se plantea las siguientes preguntas:

¿Cómo han de establecerse los términos justos de la cooperación? ¿Son acaso simplemente por alguna autoridad ajena a las personas que cooperan? ¿Acaso los establece, por ejemplo, la ley de Dios? O bien, ¿son dichas personas las que han de

²¹⁷ En adelante usaremos esta abreviatura para referirnos a la “posición original”.

reconocer como justos estos términos, tomando como referencia el saber que dispongan de un orden moral independiente?²¹⁸

La formulación de la justicia como imparcialidad implica dirigirse por la última opción señalada, buscando, también, respetar el pluralismo razonable existente en una sociedad democrática. En este sentido, Rawls reformula la doctrina del contrato social tratando de establecer las condiciones ideales para que individuos pertenecientes a una sociedad democrática –que se conciben a sí mismos como libres e iguales- establezcan un acuerdo en el que todos estén eficazmente comprometidos y todos salgan beneficiados.

El pluralismo razonable también juega un papel en la intención de cooperación, si se busca conformar una sociedad imparcialmente justa, en medio de las “circunstancias de la justicia”, que Rawls define como “condiciones históricas bajo las que existen las sociedades democráticas modernas”.²¹⁹ Dentro de estas condiciones se encuentran “las circunstancias objetivas de la justicia”, que son todas aquellas circunstancias “de escasez moderada y la necesidad de cooperación social para que todos podamos tener un nivel de vida decente”;²²⁰ asimismo, está el hecho de la inmensa diversidad de formas de concebir la vida, de las concepciones del bien diferentes unas de otras; a estas les denomina Rawls “circunstancias subjetivas de la justicia”.

Rawls hubiese podido proponer una doctrina comprensiva determinada que sostuviera toda su armazón conceptual, y a la que todos tendrían que seguir si quisieran una sociedad justa. Sin embargo, el concepto de justicia política rawlsiano no pasa por encima de

²¹⁸ John Rawls, *Liberalismo político*, Fondo de Cultura Económica, México, 1995, p. 45. Samuel Freeman, en su libro *Rawls*, al hablar de las motivaciones de Rawls para elaborar su teoría, afirma que un interés que invadió la curiosidad del filósofo estadounidense desde sus tempranos años académicos –en el marco de la Segunda Guerra Mundial- fue la realidad del mal de la humanidad frente a la benevolencia del Dios cristiano. La reflexión que hizo al respecto lo llevó a distanciarse de la doctrina cristiana: “el problema que Rawls encontró en el cristianismo no fue simplemente que Dios permitía que sucedieran maldades como el Holocausto y el bombardeo y destrucción indiscriminada de las ciudades alemanas y japonesas, y sus poblaciones civiles. Rawls cuestionaba cómo era posible la existencia de un Dios benevolente y digno de veneración que hubiera creado a la especie humana de tal modo que su voluntad fuese naturalmente corrupta y predestinada a cometer maldades, grandes y pequeñas” (Samuel Freeman, *Rawls*, Fondo de Cultura Económica, México, 2007, pp. 29-30). Por lo tanto, una de las conclusiones de Rawls fue que la moralidad no tiene necesidad de un Dios que la justifique (*Cfr.*, *Ibidem*, p. 30); en efecto, el filósofo estadounidense “recurrió a un argumento no religioso para mostrar que la justicia y el bien humano eran ‘congruentes’ y, por ende, que una sociedad plenamente justa o ‘bien ordenada’ es una ‘utopía realista’” (*Ibidem*, p. 31). Es por lo que unas de las tareas fundamentales de la cooperación social será preguntarse acerca del modo en que se construirán los términos de esta cooperación, partiendo de la suposición de que el ser humano posee disposiciones genuinas hacia la sociabilidad (*Cfr.*, *Idem*).

²¹⁹ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 123.

²²⁰ *Idem*.

la libertad de cada individuo de elegir el mejor camino que considere para vivir su vida; e, igualmente, tiene presente que todos los individuos desean llevar a cabo un estilo de vida particular, y no quiere negarle esta posibilidad a nadie.

A la apuesta por un determinado plan de vida, Rawls le denomina “identidad de intereses”, y agrega que existe un conflicto de intereses cuando los individuos buscan obtener el mayor beneficio posible encaminado a la concepción del bien que desean seguir. Esto sucede cuando no se tiene en cuenta la cooperación social. Por esta razón los principios de justicia cumplirán un papel central al momento de escoger entre diversas configuraciones sociales que determinen esta división de ventajas; a la hora de suscribir, de este modo, un acuerdo acerca de las porciones distributivas correctas.²²¹

La PO es, por consiguiente, un mecanismo de representación que modela las condiciones imparciales bajo las cuales los “representantes” de los ciudadanos deben acordar los términos equitativos de la cooperación social por los que se regirá la estructura básica de la sociedad.²²² La PO también modela las “restricciones aceptables sobre las razones”, es decir, conduce a los representantes de los ciudadanos a elegir unos principios de justicia por encima de otros.²²³

Otra posibilidad para Rawls hubiese podido ser analizar la forma en que los individuos racionales llegan a establecer determinados acuerdos a través de la toma de decisiones. Esto lo haría con información suministrada por parte tanto de la economía como de la teoría social. No obstante, esta forma de elegir diferiría significativamente de lo que se pretende en la PO; principalmente, porque, como afirma el filósofo estadounidense, “no nos proponemos describir y explicar cómo se comporta de hecho la gente en determinadas situaciones, o cómo funcionan realmente las instituciones”.²²⁴ Antes bien, con la PO se pretende “sacar a la luz una base pública para una concepción política de la justicia, hacer lo cual pertenece a la filosofía política y no a la teoría social”.²²⁵

²²¹ Cfr., *Teoría de la justicia*, *Op. cit.*, p. 126.

²²² Cfr., *La justicia como equidad*, *Op. cit.*, p. 119.

²²³ Cfr., *Idem*.

²²⁴ *Ibidem*, p. 120.

²²⁵ *Idem*.

Los representantes ideales en la PO no poseen, por ende, las características de las personas tal como son en la sociedad, en su cotidiano vivir; antes bien, “son personas artificiales, meros habitantes de nuestro mecanismo de representación: son actores que tienen un papel en la obra de nuestro experimento mental”.²²⁶ Estos individuos ideales nos ayudarán a especificar la manera en que nos imaginamos a los representantes racionales de ciudadanos libres e iguales en una democracia.²²⁷ De igual forma, estarán condicionados razonablemente por ciertas “constricciones formales” así como por un “velo de la ignorancia”, es decir, por ciertos límites a sus conocimientos.²²⁸ Se habla de condiciones “razonables”, en el sentido de que se parte de ideas intuitivas básicas que puedan “aplicarse a las personas, a sus decisiones y acciones, así como a los principios y las normas, a las doctrinas comprensivas y a otras muchas cosas”.²²⁹

Antes de adentrarnos en la explicación de las condiciones razonables mencionadas, es pertinente recordar que los representantes en la PO deberán, principalmente, abogar por el bien o, mejor dicho, por lo bienes de quienes están representando. Estos bienes, como se había señalado en el primer capítulo, los especifica Rawls como aquéllos que toda persona, concibiéndose a sí misma como libre e igual, desearía tener más que a otros. Los “bienes sociales primarios” –como los denomina Rawls- representan, a su vez, una convergencia de todas las concepciones de bien que poseen los ciudadanos libres e iguales de una sociedad democrática.²³⁰

Además, las partes en la PO tendrán que elegir dentro de una lista de principios, los cuales representan a las concepciones de justicia más importantes de nuestra tradición de filosofía política, junto con algunas otras alternativas conformadas por esta tradición.²³¹ En este sentido, los principios de justicia que elegirán los representantes ideales no son el resultado de un proceso deductivo a partir de la configuración misma de la PO, sino el

²²⁶ *Ibidem*, p. 122.

²²⁷ *Cfr.*, *Ibidem*, p. 120.

²²⁸ *Cfr.*, *Idem*.

²²⁹ *Ibidem*, p. 121.

²³⁰ Para ver la lista de los bienes sociales primarios ir a la página 22 del primer capítulo.

²³¹ *Cfr.*, *La idea de la justicia*, *Op. cit.*, p. 122. Para ver la lista completa de estos principios que menciona Rawls, *Cfr.*, *Teoría de la justicia*, *Op. cit.*, p. 124.

resultado de una lista previa. La PO es, en efecto, un mecanismo de selección que opera sobre una conocida familia de concepciones de justicia.²³²

Ahora, entonces, retomemos las condiciones razonables en la configuración de la PO. Como lo habíamos señalado, la PO nos muestra el razonamiento que Rawls establece para indicar la manera más razonable en que podríamos elegir los principios políticos de justicia que conlleven a la construcción de una sociedad justa; esto en el marco de un sistema imparcial de cooperación en un Estado democrático. Por ende, antes de explicar el camino de tal razonamiento, es importante analizar cómo se configura la PO a partir de las constricciones formales y el velo de la ignorancia.

En *Teoría de la justicia*, Rawls fija las constricciones formales del concepto de lo justo en la PO señalando que estas restricciones “sirven para escoger todos los principios éticos y no sólo aquellos de la justicia. Si las partes tuviesen que reconocer también principios para otras virtudes, estas restricciones se aplicarían igualmente”.²³³ Esto quiere decir que, si las partes quisieran favorecer cierto tipo de concepción de lo justo sobre otra, las constricciones formales se lo impedirían; como lo haría, por su función, el velo de la ignorancia, que más tarde explicaremos.

Tales constricciones rezan que los principios de justicia deben ser generales, universales y públicos. La generalidad significa que los principios “tiene que ser posible formularlos sin el uso de palabras que intuitivamente podrían ser reconocidas como nombres propios o descripciones definidas”.²³⁴ Esta condición se desprende del hecho de que las partes no poseen información de sí mismas; por esta razón, no podrían formular principios expresando cierta autoidentificación en los mismos. El hecho de que los principios sean generales también radica en la necesidad de que estos sean capaces de servir como base pública perpetua de una sociedad bien ordenada.²³⁵

La universalidad, por su parte, se refiere al hecho de que los principios puedan ser “aplicados a todos los agentes morales sin inconsistencia o incoherencia contraproducente,

²³² Cfr., *La idea de la justicia*, *Op. cit.*, p. 122.

²³³ John Rawls, *Teoría de la justicia*, *Op. cit.*, 129.

²³⁴ *Ibidem*, p. 130.

²³⁵ Cfr., *Ibidem*, p. 131.

en nuestro caso, a todos los ciudadanos de la sociedad en cuestión”.²³⁶ En efecto, los dos principios de justicia son los que, para Rawls, cumplen con tal restricción. Decantarse por tales principios universales desde un ideal democrático, a partir de una lista previa procedente de la tradición de la filosofía política, ha generado algunas críticas a Rawls en las que se afirma, por ejemplo, que en su teoría se da una “manipulación de cartas a su favor, haciendo suposiciones en la PO que solamente conducen a sus principios de justicia”.²³⁷ Sin embargo, Samuel Freeman cuestiona tales críticas planteándose la siguiente interrogante:

¿por qué habría de ser esto un defecto y no una virtud de la teoría, en tanto que estas suposiciones no son arbitrarias y realmente capturan mejor que cualquiera otra opción nuestras convicciones consideradas acerca de restricciones razonables sobre argumentos en favor de principios de justicia?²³⁸

Aunque nuestra intención no es profundizar en este tipo de análisis, sí es pertinente anotar que la mayor parte de críticas que se le hacen a Rawls ignoran las bases argumentativas que el autor establece para cada una de sus ideas principales. Sobre esta idea volveremos, principalmente, cuando problematicemos las críticas de Harsanyi a Rawls. Por ahora, sigamos con otra de las restricciones formales.

La última de las restricciones formales al concepto de lo justo que señala Rawls es que los principios de justicia deben ser públicos. La “publicidad” amerita, para el profesor de Harvard, que “las partes en la posición original, al evaluar los principios, deban tener en cuenta las consecuencias, sociales y psicológicas, del reconocimiento público por parte de los ciudadanos de que dichos principios son mutuamente aceptados y regulan efectivamente la estructura básica”.²³⁹ Rawls distingue tres niveles de publicidad que una sociedad bien ordenada debe conseguir: “el reconocimiento mutuo, por parte de los ciudadanos, de los principios de justicia junto con el conocimiento público de que las instituciones de la estructura básica realmente satisfacen dichos principios”;²⁴⁰ el segundo nivel requiere un “reconocimiento mutuo, por parte de los ciudadanos, de los hechos generales, sobre la base de los cuales las partes en la PO seleccionan esos principios”;²⁴¹ y el tercer nivel, de

²³⁶ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 125.

²³⁷ Samuel Freeman, *Rawls*, *Op. cit.*, p. 144.

²³⁸ *Idem.*

²³⁹ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 126.

²⁴⁰ *Ibidem*, p. 166.

²⁴¹ *Idem.*

publicidad, insta a un “reconocimiento mutuo de la completa justificación de la justicia como equidad en sus propios términos”.²⁴²

En *Teoría de la justicia*, Rawls señala también dos condiciones más que no puntualiza en *La justicia como equidad*. Una primera condición es la “ordenación de las demandas conflictivas”, esta se refiere a que una concepción de justicia debe ser completa y debe ser capaz de “ordenar todas las demandas que puedan surgir (o que es probable que surjan en la práctica) [...] tal ordenación debería ser transitiva: si, digamos, una primera configuración de la estructura básica es calificada como más justa que otra segunda, y la segunda como más justa que otra tercera, entonces la primera deberá ser más justa que la tercera”.²⁴³ La segunda condición trata del “carácter definitivo” de los principios de justicia, el cual consiste en que “las partes han de considerar el sistema de principios como tribunal supremo de apelación en materia de razonamiento público; no existen normas más elevadas a las que puedan remitirse los argumentos que apoyen demandas; un buen razonamiento a partir de esos principios es concluyente”.²⁴⁴

Rawls sintetiza las cinco restricciones formales señaladas de la siguiente forma: “una concepción de lo justo es un conjunto de principios, de forma general y universales en su aplicación, que han de ser públicamente reconocidos como tribunal final de apelación para jerarquizar las demandas conflictivas de las personas morales”.²⁴⁵ Además, agrega que estas restricciones ayudarán a desechar distintas formas de egoísmo, tales como: “el egoísmo en la forma de dictadura unipersonal”,²⁴⁶ el cual es el propio de una forma de gobierno dictatorial. Otro tipo de egoísmo que enfrentan estas constricciones es el hecho de que los principios de justicia sirvan a un pequeño grupo y no a todos los ciudadanos; sin embargo, en la condición de generalidad no se excluye el egoísmo general, es decir, la posibilidad de que cada persona posea una concepción del bien que elija para desenvolverse a lo largo de su vida. La condición de “ordenación jerárquica” ayuda, a su vez, a que este egoísmo general no desemboque en una tiranía o sistema absolutista basado en la astucia o la fuerza.²⁴⁷ Para

²⁴² *Idem*.

²⁴³ John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 133.

²⁴⁴ *Ibidem*, pp. 133-134.

²⁴⁵ *Ibidem*, p. 134.

²⁴⁶ *Ibidem*, p. 131.

²⁴⁷ *Cfr.*, *Ibidem*, p. 134.

Rawls, entonces, una sociedad equitativamente justa debe llevar a establecer una serie de deberes y obligaciones que “son la consecuencia de unos primeros principios que valen para todos en tanto que personas morales; la derivación de tales exigencias tiene una base común”.²⁴⁸

I. Velo de la ignorancia: dispositivo de imparcialidad

Además de las restricciones formales, existe otro elemento fundamental en la configuración de la PO; estamos hablando del velo de la ignorancia. Este dispositivo ayuda a establecer un procedimiento imparcialmente justo en la elección de los principios. Bajo el velo de la ignorancia las partes “no saben cómo las diversas alternativas afectarán sus propios casos particulares, viéndose así obligadas a evaluar los principios únicamente sobre la base de consideraciones generales”.²⁴⁹ Específicamente, tras este velo,

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 [...] 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 [...] No tienen ninguna información respecto a qué generación pertenecen.²⁵⁰

De esta forma, los únicos hechos que conocen los representantes ideales en la PO son los relativos a las “circunstancias de la justicia”, (las cuales describimos antes); además, conocen los hechos generales acerca de la sociedad humana:

[...]entienden las cuestiones políticas y los principios de la teoría económica; conocen las bases de la organización social y las leyes de la psicología humana [...] Se supone que conocen todos los hechos generales que afectan la elección de los principios de justicia. No existen limitaciones a la información general, esto es, acerca de las teorías y leyes generales, ya que las concepciones de la justicia tienen

²⁴⁸ *Ibidem*, p. 132.

²⁴⁹ *Ibidem*, p. 135.

²⁵⁰ *Ibidem*, pp. 135-136.

que ajustarse a las características de los sistemas de cooperación social que han de regular, por lo cual no hay razón para eliminar estos hechos.²⁵¹

Las principales tradiciones de filosofía política serán el sustrato desde el cual las partes tendrán que elegir, a partir de una lista previa, los principios de justicia; esto, necesariamente, los lleva a tener a su disposición esos hechos generales a los cuales se refiere Rawls. Además, hay que tener presente la diferencia entre el conocimiento que podrían poseer las partes de ciertos hechos particulares de la sociedad o, en efecto, de la situación política de su sociedad –a los que no podrían tener acceso tras el velo de la ignorancia-, y el conocimiento general que estas tienen de las cuestiones políticas y de la teoría económica estando tras el velo de la ignorancia. Para el filósofo de Harvard, es importante que los representantes ideales conozcan estos hechos generales, ya que será fundamental que, luego de promulgar los principios de justicia, los ciudadanos muestren la disposición de apoyar tal resolución; esto conllevaría, de igual forma, a asegurar la estabilidad de la cooperación social.

Una particularidad que tiene la PO como mecanismo de representación es que “una o más personas pueden entrar en cualquier momento en esta posición, acaso mejor, simular las deliberaciones de esta situación hipotética, razonando simplemente conforme a las restricciones apropiadas”.²⁵² Sin las constricciones formales al concepto de lo justo y las restricciones al conocimiento, tal como se han planteado, es probable que se desvirtúe la finalidad para la cual ha sido creada la PO. Además, no se debe relacionar este escenario de elección con un tiempo o grupo de personas específicos; antes bien, “la PO tiene que ser interpretada de manera que en cualquier momento se pueda adoptar su perspectiva”.²⁵³ Por lo tanto, según Rawls, siguiendo las pautas propuestas dentro del planteamiento de la PO, los sujetos ideales siempre acordarán la elección de los mismos principios de justicia; el velo de la ignorancia será una pieza fundamental para tal fin.²⁵⁴

²⁵¹ *Ibidem*, p. 136.

²⁵² *Ibidem*, pp. 136-137.

²⁵³ *Ibidem*, p. 137.

²⁵⁴ Al parecer, el velo de la ignorancia es una de las ideas más originales e importantes de Rawls; no obstante, para Brian Barry, “el ‘velo de la ignorancia’ en sí mismo no es lo distintivo en el enfoque de Rawls. Es compatible con cualquier concepto de la moralidad que insista en el papel central de la imparcialidad” (Brian Barry, *La teoría liberal de la justicia: examen crítico de las principales doctrinas de Teoría de la justicia de John Rawls*, Fondo de Cultura Económica, México, 1993, p. 23). Uno de los ejemplos que presenta Barry para sustentar esta premisa es la propuesta de R.M Hare; específicamente, las limitaciones que impone al contenido y a las formas posibles de los principios morales (*Cfr.*, *Ibidem*, p. 22.). Según Barry, tales limitaciones tendrían

Hasta aquí hemos explicado la configuración de la PO, es decir, los distintos aspectos que llevan a crear un escenario propicio para que los representantes ideales escojan principios de justicia que identifiquen a una sociedad equitativamente justa. Ahora veremos el tipo de razonamiento que sigue Rawls a partir de la naturaleza de la PO para, así, sustentar el por qué sus dos principios de justicia política serían los más racionales y razonables para una sociedad democrática con régimen constitucional. Esto nos ayudará también a retomar y profundizar las principales críticas de Rawls a la tradición utilitarista, en vistas de entender las principales razones por las que, para el filósofo estadounidense, no sería racional ni razonable elegir el principio de utilidad en la PO.

2. Principios de justicia: una alternativa a la propuesta utilitarista

Antes de explicar las principales razones por las que los representantes ideales en la PO eligen los principios de justicia sobre el principio de utilidad será pertinente exponer, *grosso modo*, el planteamiento de los dos principios que propone Rawls a partir de los “bienes sociales primarios” que ha formulado.

Un individuo en la PO, teniendo en cuenta la configuración de ésta, le será imposible tratar de obtener ventajas sólo para él mismo; tampoco habría manera de que le correspondiera aceptar desventajas. Por lo tanto, “lo más sensato es reconocer como primer paso un principio de justicia que exija una distribución igualitaria”.²⁵⁵ En efecto, la primera resolución de las partes en la PO será optar por una igualdad de libertades básicas para todos, una igualdad equitativa de oportunidades y una división igualitaria de ingresos y riquezas.²⁵⁶

Ahora, haciendo uso de los hechos generales a los cuales tienen acceso en la PO, los representantes saben que la eficiencia económica es importante en una sociedad, así como la exigencia de la organización y la tecnología. Entonces, un reto que se les presenta es encontrar el mejor escenario para que la situación de igualdad total inicial no empeore, sino que mejore haciendo uso –de alguna forma- de los bienes sociales primarios. Las partes, entonces, decidirán que las desigualdades en el bien primario de ingresos y de riqueza

el mismo objetivo que el velo de la ignorancia. Además, a partir de la curiosa similitud entre la obra *Freedom and Reason*, de Hare, y la *Teoría de la justicia*, Barry concluye que los dos autores, al parecer, estaban trabajando en la misma cuestión, y sin embargo, no fueron lo suficientemente conscientes de ello.

²⁵⁵ John Rawls, *La teoría de la justicia*, *Op. cit.*, pp. 147-148.

²⁵⁶ *Cfr.*, *Ibidem*.

ayudarán a mejorar la situación de todos, incluyendo a los menos aventajados;²⁵⁷ y estos últimos, a su vez, tienen un derecho de veto; de esta manera se llega al principio de diferencia, es decir: “tomando la igualdad como punto de comparación, aquellos que han ganado más tienen que haberlo hecho en términos que sean justificables respecto a aquellos que han ganado menos”.²⁵⁸ De esta manera, los principios de justicia quedan formulados de la siguiente forma:

- a) Cada persona tiene el mismo derecho irrevocable a un esquema plenamente adecuado de libertades básicas iguales que sea compatible con un esquema similar de libertades para todos; y
- b) Las desigualdades sociales y económicas tienen que satisfacer dos condiciones: en primer lugar, tienen que estar vinculadas a cargos y posiciones abiertos a todos en condiciones de igualdad equitativa de oportunidades; y, en segundo lugar, las desigualdades deben redundar en un mayor beneficio de los miembros menos aventajados de la sociedad (el principio de diferencia).²⁵⁹

Adentrémonos ahora, entonces, en el razonamiento que establece Rawls en la PO, el cual se organiza a partir de dos comparaciones fundamentales.²⁶⁰ Estas comparaciones ayudan a

²⁵⁷ Cabe destacar que Rawls, para elegir el principio de diferencia, resaltó la importancia de la igualdad democrática por encima de la aristocracia natural, el sistema de libertad natural y la interpretación liberal. La igualdad democrática intenta hacerle frente a la inevitable influencia de los efectos arbitrarios de la lotería natural. Es por ello que el filósofo estadounidense propone el principio de diferencia por encima del principio de eficiencia que resultaría de una interpretación liberal, donde la distribución del ingreso y la riqueza quedaría a merced de la manera en que cada individuo ha desarrollado sus talentos naturales, a partir de la garantía de la defensa de sus libertades básicas y una justa igualdad de oportunidades. El principio de diferencia, entonces, ayudará a que las expectativas más elevadas de quienes están mejor situados sean justas si y sólo si funcionan como parte de un esquema que mejore las expectativas de los menos favorecidos de la sociedad. Sin embargo, aunque el principio de diferencia ayude a mitigar el efecto de la distribución de las capacidades innatas y las contingencias que le ha tocado vivir a cada individuo, es innegable que la lotería natural seguirá jugando un papel fundamental en la cooperación productiva de la sociedad. Aun así, el principio de diferencia es el mejor escenario posible que puede acercar a la sociedad a un punto donde se estrechen, cada vez más, las desigualdades en ingreso y riquezas. Cabe señalar, en efecto, que el principio de diferencia se inspira en la idea de reciprocidad que juega un papel decisivo al momento de confrontar los principios rawlsianos con el principio utilitarista. Más adelante profundizaremos en este aspecto.

²⁵⁸ *Ibidem*, p. 148.

²⁵⁹ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 73. La razón por la cual Rawls establece un escenario de desigualdad en los bienes sociales primarios de ingresos y riquezas no es muy clara en los argumentos que presenta en *La teoría de la justicia*. Los trabajos que hace después (como en *La justicia como equidad: una reformulación* y en *Liberalismo político*), tratando de corregir o aclarar puntos relevantes de su teoría, dan un poco más de claridad acerca del principal argumento que sustenta tal desigualdad. No obstante, nuestro objetivo en este trabajo no se centra en la búsqueda de tal argumentación; antes bien, deseamos –por ahora– conocer la forma en que se plantean los dos principios de justicia que los representantes ideales contrastarán con el principio de utilidad dentro del razonamiento que hacen en la PO.

²⁶⁰ Esta forma de organizar el razonamiento en la PO lo establece Rawls años después de haber publicado *A Theory of Justice*, específicamente, en el texto publicado en 1974 titulado *Reply to Alexander and Musgrave*. Esta misma formulación del razonamiento es la que se encuentra en *La justicia como equidad*, la cual nos servirá de guía para seguir nuestra argumentación.

comprender las razones que llevan al filósofo estadounidense a proponer cada uno de los principios de justicia. Específicamente, tales comparaciones coadyuvan a entender el “principio de diferencia en tanto principio de justicia distributiva y la regla maximin como regla práctica para decisión bajo incertidumbre”.²⁶¹ Esto ayudará a evitar caer en confusiones respecto de la función de la regla maximin en el conjunto de la teoría de la justicia rawlsiana, y de igual forma, ayudará a precisar, un poco más, la distancia que toma Rawls del utilitarismo.

Rawls recurrió a la herramienta de las comparaciones fundamentales también para brindar más claridad en cuanto a un argumento razonablemente concluyente en favor de los dos principios de justicia.²⁶² Esto finalmente nos conducirá, por lo tanto, a profundizar en las principales críticas al utilitarismo; especialmente, al utilitarismo contemporáneo de Harsanyi. Específicamente, aquí analizaremos las implicaciones que traería a una sociedad democrática el hecho de que los representantes ideales elijan al principio de utilidad como principio político para una sociedad que pretende ser equitativamente justa.

Rawls comparará los dos principios de justicia con las alternativas que presenta la lista previa de concepciones políticas. En este sentido, el filósofo estadounidense buscará un balance más robusto de razones en cada comparación, lo que ayudará a ratificar el por qué los dos principios de justicia serían los más apropiados. En efecto, los dos principios de justicia serán comparados, “de forma diferente cada vez, con el principio de utilidad media, y la comparación muestra, en el mejor de los casos, su superioridad sobre dicho principio”.²⁶³ Esta comparación nace de dos ideas de sociedad que han ocupado un lugar importante en la tradición del pensamiento democrático. Una de las ideas se refiere:

A la sociedad como un sistema equitativo de cooperación social entre ciudadanos concebidos como libres e iguales; la otra es la idea de la sociedad como un sistema social organizado de tal modo que produzca el mayor bien agregado a todos sus miembros, donde ese bien es un bien completo determinado por una doctrina comprensiva.²⁶⁴

En otras palabras, se trata de la tradición del contrato social y la tradición utilitarista. La tradición contractualista, caracterizada por la búsqueda de una sociedad equitativamente

²⁶¹ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 136.

²⁶² *Cfr.*, *Ibidem*, p. 136.

²⁶³ *Ibidem*, p. 136.

²⁶⁴ *Ibidem*, p. 137.

cooperativa, tiene como ejes centrales tanto la igualdad como la reciprocidad. Dentro del primer eje se ubican tanto los derechos y libertades básicos, como las oportunidades equitativas; el segundo eje, la reciprocidad, se centra en el principio de diferencia. La tradición utilitarista, por su parte, instrumentaliza tanto la igualdad como la reciprocidad buscando maximizar la suma del bienestar social.²⁶⁵ De esta manera, las dos comparaciones fundamentales que presenta Rawls giran en torno a esta marcada diferencia. En la primera comparación, entonces, se mostrará por qué los dos principios de justicia son superiores al principio de utilidad con respecto a la igualdad. Y en la segunda comparación se analizará la ventaja de los principios de justicia sobre el principio de utilidad desde la idea de reciprocidad.²⁶⁶

I. La regla maximin como directriz: primera comparación fundamental

En la primera comparación fundamental se contrastan los dos principios de justicia –tomados como una unidad- con el principio de utilidad media. Teniendo en cuenta el planteamiento de Harsanyi, la organización de una sociedad, liderada por el principio de utilidad promedio, estaría conformada por las instituciones de la estructura básica dispuestas de tal modo “que maximicen el bienestar medio de los miembros de la sociedad, comenzando en el momento presente y extendiéndose hasta el futuro previsible”.²⁶⁷ El principal argumento de Rawls para confrontar el principio de utilidad –desde el planteamiento de Harsanyi- en la primera comparación fundamental es la regla maximin, la cual “nos dice que identifiquemos al peor resultado de cada alternativa disponible y que adoptemos, seguidamente, la alternativa cuyo peor resultado es mejor que los peores resultados de todas las demás alternativas”.²⁶⁸

La regla maximin persigue uno de los principales objetivos de la justicia como equidad, esto es, encontrar una base moral más apropiada –que el utilitarismo, principalmente- para las instituciones de una sociedad democrática moderna.²⁶⁹ Por lo tanto,

²⁶⁵ *Cfr.*, *Ibidem*, p. 137.

²⁶⁶ *Cfr.*, *Idem*.

²⁶⁷ *Idem*.

²⁶⁸ *Ibidem*, p. 138.

²⁶⁹ *Cfr.*, *Idem*. Recordemos que Rawls elabora su propuesta de justicia tratando de hallar una alternativa a concepciones que han significado mucho en la tradición moral y política; estamos hablando del utilitarismo, perfeccionismo y el intuicionismo. No obstante, como habíamos señalado, el profesor de Harvard se centrará en una fuerte crítica al utilitarismo por considerarla la concepción que ha tenido más relevancia en las sociedades democráticas modernas.

la regla maximin ayudará a que los representantes ideales tengan como referencia las peores posiciones sociales en una sociedad dirigida por los principios de justicia. Rawls expone el argumento de la siguiente manera:

- i) Si hay ciertas condiciones en las que es racional guiarse por la regla maximin cuando se acuerdan principios de justicia para la estructura básica, entonces, en esas condiciones, se acordarían los dos principios de justicia en vez del principio de utilidad media.
- ii) Hay ciertas condiciones, tres en particular, tales que, cuando se cumplen, es racional guiarse por la regla maximin cuando tratamos de acordar principios de justicia para la estructura básica.
- iii) Esas tres condiciones se cumplen en la posición original.
- iv) Por consiguiente, los dos principios de justicia serían acordados por las partes en vez del principio de utilidad media.²⁷⁰

Las tres condiciones que señala Rawls en el segundo punto del argumento planteado son las siguientes: la primera condición consiste en que “las partes no tienen ningún modo fiable de estimar las probabilidades de las posibles circunstancias sociales que afectan a los intereses fundamentales de las personas a las que representan”.²⁷¹ Es decir, a diferencia del modelo de equiprobabilidad de Harsanyi como punto de partida de su propuesta ética, Rawls descarta un razonamiento que parta del concepto de probabilidad. Para Harsanyi, un juicio de valor genuinamente moral es aquél donde el observador imparcial elige asumiendo que, en cualquier tipo de sociedad por la que haya optado, tendrá la misma posibilidad de ocupar cualquier tipo de posición social, desde la menos aventajada hasta la más aventajada. No obstante, para Rawls el sentido de justicia que poseen los representantes ideales es una facultad moral que –desde la regla maximin- los lleva a decidir evaluando las alternativas desde sus peores resultados posibles. Esto quiere decir que:

debe ser racional para las partes [...] no prestar demasiada atención a lo que podría ganarse más allá de lo que puede ser garantizado (para aquellos a los que representan) adoptando la alternativa cuyo peor resultado es mejor que los peores resultados de todas las demás alternativas.²⁷²

A la alternativa cuyo peor resultado es el mejor le denomina Rawls “nivel garantizable”.²⁷³ En este sentido, la segunda condición, dentro del argumento de la primera comparación,

²⁷⁰ *Ibidem*, pp. 138-139.

²⁷¹ *Ibidem*, p. 140.

²⁷² *Idem*.

²⁷³ *Idem*.

amerita que el nivel garantizable posea un nivel suficiente de satisfacción. Y la tercera condición insta a que los peores resultados de todas las demás alternativas estén significativamente por debajo del nivel garantizable.²⁷⁴ Para Rawls, el uso de estas condiciones obedece a un principio de maximización de la satisfacción del bien racional. Las partes, por medio de la regla maximin, se sirven de tal principio de maximización para “guiarse en las circunstancias extremadamente inusuales, si no únicas, de la posición original”.²⁷⁵

La regla maximin tiene un valor práctico al momento de establecer qué tipos de intereses se podrían considerar fundamentales en la conformación de los principios de justicia para la estructura básica de la sociedad. Incluso, desde el argumento maximin, Rawls añade que las partes tienen como referencia la maximización de la utilidad esperada. Sin embargo, tal maximización no se hace desde un contenido sustantivo, como lo harían autores como Sidgwick o Harsanyi. Es decir, los agentes racionales en la PO se valen de la utilidad esperada en tanto un orden o clasificación “en que las alternativas se juzguen mejores o peores en punto de satisfacer los intereses fundamentales de los agentes, que en este caso son los intereses de ciudadanos libres e iguales”.²⁷⁶

En la PO se busca, entonces, establecer un nivel garantizable que, a su vez, no esté desprovisto de una maximización de la utilidad esperada de los agentes. En la regla maximin, en efecto, el nivel garantizable se traduce en esto: “la situación de los miembros menos aventajados de la sociedad bien ordenada que resulta de la plena realización de los principios de justicia”.²⁷⁷

Las tres condiciones, por ende, nos ayudarán a demostrar porqué el principio de utilidad media no sería el más favorable en la elección de las partes en la PO. Antes bien, tales condiciones demostrarían por qué los dos principios de justicia serían los más adecuados para la estructura básica de una sociedad. Específicamente, partiendo del hecho de que los representantes ideales descartan las estimaciones de probabilidad –primera condición–, “el conocimiento y las creencias bien fundadas sobre probabilidades deben basarse, cuando menos, en ciertos hechos establecidos o en creencias sobre el mundo bien respaldadas

²⁷⁴ *Cfr., Idem.*

²⁷⁵ *Idem.*

²⁷⁶ *Ibidem*, pp. 140-141.

²⁷⁷ *Ibidem*, p. 141.

empíricamente”²⁷⁸. Rawls afirma, además, que los representantes conocen la existencia de las “circunstancias de la justicia” bajo condiciones razonablemente favorables; para que se den tales condiciones debe existir, no obstante, voluntad política que, a su vez, hará posible un régimen constitucional. Sin embargo, la existencia de las condiciones favorables no es suficiente para que las partes “puedan establecer una distribución bien fundada de probabilidades sobre las formas de cultura y tradición política que pudieran darse”.²⁷⁹

Sumado a lo anterior, Rawls hace una distinción entre “riesgo” e “incertidumbre”, señalando que el riesgo posee cierta parte de evidencia objetiva para estimar las probabilidades,²⁸⁰ mientras que en la incertidumbre no existe tal base objetiva, ya que si existiese tal base sería totalmente imprecisa. Por lo tanto, los representantes, en la PO, se encuentran en situación de incertidumbre más que en situación de riesgo, es decir, “no poseen una base fiable para estimar las probabilidades de las posibles condiciones sociales e históricas, o la probabilidad de que las personas a quienes representan profesen una doctrina comprensiva (con su concepción del bien) y no otra”.²⁸¹

Una de las críticas que se le hacen a Rawls en este aspecto es que parece que los representantes, al hacer uso de la regla maximin, poseen una irracional u obsesiva aversión a la incertidumbre. No obstante, el filósofo estadounidense responde que los representantes, como fideicomisarios de los intereses de los ciudadanos, deberán velar por los intereses fundamentales de sus representados. En efecto, la regla maximin sirve para tal objetivo. Por ejemplo, al no saber qué posición le corresponderá a un ciudadano en la sociedad, lo más racional para las partes será asegurar que a las personas que queden en la posición menos aventajada se les garantice la defensa de sus derechos y libertades básicos, oportunidades equitativas, así como el acceso a “una parte adecuada de los medios materiales de uso universal (los bienes primarios del ingreso y la riqueza)”;²⁸² de tal modo que los ciudadanos representados, aun quedando en la posición menos aventajada, sean capaces de ejercer esos derechos y libertades y de sacar provecho de esas oportunidades.²⁸³

²⁷⁸ *Ibidem*, p. 143.

²⁷⁹ *Ibidem*, p. 143.

²⁸⁰ *Cfr.*, *Ibidem*, p. 148.

²⁸¹ *Ibidem*, 149.

²⁸² *Idem*.

²⁸³ *Cfr.*, *Idem*.

Ahora, el cumplimiento de la segunda y tercera condición son un argumento poderoso a favor de los dos principios de justicia. Profundicemos en estas dos condiciones para analizar, de una forma más detallada, por qué no sería viable elegir el principio de utilidad en la PO. Empecemos por el hecho de que el nivel garantizable altamente satisfactorio, que demanda la segunda condición, está asegurado por la defensa de los derechos y libertades básicos e iguales para todos en los dos principios de justicia; mientras que el principio de utilidad “puede a veces permitir o, más aún, exigir la restricción o supresión de los derechos y libertades de algunos en aras de un agregado mayor de bienestar social”.²⁸⁴ Es por esto que lo más razonable para las partes es abogar por la defensa de los intereses fundamentales de sus representados, lo cual, en efecto, los llevaría a elegir los dos principios de justicia por encima del principio de utilidad media.

El velo de la ignorancia deja a los representantes desprovistos de la posibilidad de conocer si las personas a las que representan pertenecen a un determinado grupo religioso, político o de cualquier otra índole. Esta es una razón por la cual las partes no podrían dejar a un lado la defensa de los intereses fundamentales de sus representados. Los representantes, por ejemplo, “no pueden correr riesgos permitiendo una menor libertad de conciencia para las minorías religiosas”.²⁸⁵ De esta manera, un ciudadano tendrá asegurada su libertad de conciencia no importando su pertenencia a una religión dominante o minoritaria. A la defensa de las libertades básicas como intereses fundamentales la denomina Rawls “primacía de la libertad”, y la cual significa que:

Una libertad básica sólo puede ser limitada o conculcada en aras de una o más de las otras libertades básicas, y nunca en aras de un bien público mayor, entendido como una suma neta mayor de ventajas sociales y económicas para la sociedad en su conjunto.²⁸⁶

Fijar la primacía de los derechos y libertades básicos, en una sociedad democrática bajo un régimen constitucional, ayuda a que, en medio del pluralismo razonable, las discusiones acerca de los derechos y libertades básicos no queden a merced de intereses mezquinos de partidos políticos y del proceder insensato de representantes de las instituciones políticas y sociales. Además, esto lleva a establecer que los términos de la cooperación social deban

²⁸⁴ *Ibidem*, p. 144.

²⁸⁵ *Ibidem*, p. 147.

²⁸⁶ *Ibidem*, p. 154.

descansar sobre la base del respeto mutuo.²⁸⁷ Si hablamos de una sociedad democrática dirigida por el principio de utilidad media, los derechos y libertades básicos penderían del cálculo de los intereses sociales, buscando, principalmente, la maximización del bienestar del mayor número de ciudadanos. Este hecho, según Rawls, dejaría a una sociedad desprovista de unos intereses fundamentales igualitarios dignos de ser protegidos y respetados por todos los que hacen parte de dicha sociedad. Esto lleva a afirmar, de igual manera, que “los dos principios de justicia establecen una base más clara y fiable de razón pública que el principio de utilidad”.²⁸⁸

Hasta aquí hemos analizado, entonces, la forma en que proceden los representantes en la PO a partir de la regla maximin. Para ello, hemos señalado tres condiciones que se desprenden de tal regla y que se presentan como razón suficiente para elegir los dos principios de justicia sobre el principio de utilidad media. Uno de los argumentos que lleva a los representantes a inclinarse por los dos principios de justicia se desprende de la tercera condición. En tal argumento se sostiene que es indispensable la defensa de los intereses fundamentales de los ciudadanos; específicamente, estamos hablando de los derechos y libertades básicos, sin los cuales sería difícil mantener una base de razón pública en una sociedad con régimen constitucional, pues podrían producirse transgresiones en contra del respeto mutuo y el pluralismo razonable.

Además, existe otro argumento a favor de los dos principios de justicia que se desprende de la segunda condición en el planteamiento de la regla maximin. Este argumento va dirigido, principalmente, al hecho de cómo establecer un nivel garantizable satisfactorio. De este modo, partiendo de la primacía de los derechos y libertades básicos, la sociedad debe buscar la manera de ayudar a consolidar y fortalecer una base de razón pública clara y fiable para todos los ciudadanos, lo cual no podría lograr el principio de utilidad media. Para consolidar una base de razón pública clara se deben, también, estimular las virtudes cooperativas de la vida política, que Rawls distingue como “las virtudes de la razonabilidad y del sentido de la equidad, del espíritu de compromiso y de la disposición a asumir compromisos con los demás”.²⁸⁹ Tales virtudes ayudarán a acrecentar el deseo de cooperar

²⁸⁷ *Cfr.*, *Ibidem*, p. 159.

²⁸⁸ *Ibidem*, p. 160.

²⁸⁹ *Ibidem*, p. 161.

con los otros, “en términos que todos puedan aceptar públicamente como equitativos sobre una base de igualdad y respeto mutuo”.²⁹⁰

Rawls se vale de la regla maximin para hacer una primera comparación entre los dos principios de justicia, tomándolos como una unidad, y el principio de utilidad. Hace esta primera comparación fundamental resaltando la pertinencia y el valor de los derechos y libertades básicos en una sociedad democrática desde el ideal de igualdad. En una segunda comparación, Rawls contrasta los dos principios de justicia con el principio de utilidad restringida; en ésta él resaltará la importancia del principio de diferencia frente a la propuesta utilitarista desde la idea de reciprocidad. Veamos la manera en que procede el filósofo estadounidense para llevar a cabo esta segunda comparación fundamental.

II. Principio de diferencia versus principio de utilidad restringida: segunda comparación fundamental

En *Teoría de la justicia*, al tratar de ofrecer una alternativa a la clásica doctrina utilitarista, Rawls establece comparaciones entre los principios de justicia frente a ciertas concepciones mixtas de los mismos,²⁹¹ señalando que todas estas aceptarían el primer principio y, por lo tanto, reconocerían la primacía de la libertad. En este sentido, a una de las concepciones mixtas que el filósofo estadounidense encuentra más cercana a los principios de justicia, la denomina “principio de utilidad restringida”. En esta los dos principios de justicia, tomados como una unidad, se comparan con dos principios similares que, en vez del principio de diferencia, están dirigidos por el principio de utilidad media. Es decir, aquí la estructura básica estaría organizada de tal manera que ayudaría a maximizar la utilidad media garantizando, también y en primer lugar, las libertades básicas iguales y la equitativa igualdad de oportunidades. La única diferencia, pues, consiste en que este principio de utilidad media debe mantener un “mínimo social apropiado”.²⁹² ¿Será ésta la vía más efectiva para hablar de una PO utilitarista? Rawls responderá que el principio de utilidad restringida es insuficiente para poder hablar de una sociedad equitativamente justa. Y la principal razón de

²⁹⁰ *Idem.*

²⁹¹ *Cfr., Teoría de la justicia, Op. cit., pp. 291-299.*

²⁹² John Rawls, *La justicia como equidad, Op. cit., p. 165.*

esta conclusión se haya en la idea de “reciprocidad”, la cual se halla implícita en el principio de diferencia y de la cual carece el principio de utilidad restringida.

La razón más relevante que distancia al principio de utilidad restringida del principio de diferencia es, en efecto, esta idea de reciprocidad, que Rawls desarrolla más claramente en *La justicia como equidad*.²⁹³ En la primera comparación se destaca la importancia de la igualdad al momento de determinar entre qué principios elegir. Esta segunda comparación muestra, por su parte, la ventaja de los dos principios de justicia sobre el principio de utilidad restringida desde la noción de reciprocidad.

Para explicar el sentido y lugar de la reciprocidad en una sociedad equitativamente justa, debemos recordar que tal sociedad está conformada por dos grupos de personas: el grupo de los más aventajados y el grupo de los menos aventajados. Uno de los criterios para este tipo de clasificación es la desigualdad existente entre los bienes primarios de ingresos y de riqueza. Como habíamos señalado antes, el principio de diferencia regula estas desigualdades. Para el planteamiento del principio de diferencia, Rawls procede con el siguiente hilo argumentativo:

Puesto que las partes en la posición original están situadas simétricamente y saben (por el contenido común de las dos alternativas) que el principio adoptado se aplicará a ciudadanos concebidos como libres e iguales, arrancan de la igual división de ingresos y riquezas (iguales perspectivas de vida tal cual quedan indexadas por los bienes primarios) como punto de partida. Entonces se preguntan: ¿hay buenas razones para apartarse de la división igual y, de haberlas, qué desigualdades, surgidas de qué modo, son aceptables?²⁹⁴

Los representantes, entonces, eligen un escenario de desigualdad siempre y cuando, desde un punto inicial de igualdad, se mejore la situación de todos los ciudadanos. Por consiguiente, el principio de diferencia es el resultado de tomar la igual división como punto de partida e introducir una idea de reciprocidad.²⁹⁵ Esta idea “selecciona un punto focal natural entre las demandas de la eficiencia y la igualdad”.²⁹⁶ La reciprocidad implica que “las instituciones sociales no deben sacar provecho de las contingencias de las dotaciones innatas, o de la posición social, o de la buena o mala suerte durante el transcurso de la vida, a no ser que ello

²⁹³ Cfr., *Ibidem*, pp. 164-173.

²⁹⁴ *Idem*.

²⁹⁵ Cfr., *Ibidem*, p. 169.

²⁹⁶ *Idem*.

beneficie a todos, incluidos los menos favorecidos”;²⁹⁷ esto debido a que en la idea de reciprocidad se encuentra implícita la consideración de la distribución de las dotaciones innatas como un activo común.²⁹⁸

Dentro del concepto de reciprocidad, que se desprende del principio de diferencia, están implícitas las ideas de publicidad y estabilidad. La publicidad, como lo habíamos mencionado, exige que las partes evalúen los principios de justicia a la luz de las consecuencias –políticas, sociales y psicológicas– del reconocimiento público, por parte de los ciudadanos en general, de que dichos principios serían aprobados por ellos y que regularían efectivamente la estructura básica; mientras que la estabilidad ayuda a constituir una sociedad donde “los ciudadanos aceptan como justas las instituciones existentes y no suelen tener deseo alguno de violar o renegociar los términos de la cooperación social, dada su posición social presente y prospectiva”.²⁹⁹

Una de las amenazas a la “estabilidad sociopolítica” es actuar de tal manera que se trate de perseguir intereses personales sin importar cómo afectará tal proceder a los demás integrantes de la sociedad. Sumado a esto, está el hecho del descontento por parte de los menos aventajados debido a su posición en la sociedad; en este sentido, es inevitable que los menos aventajados sean “los más prestos a violar los términos de la cooperación o a urgir a una renegociación”.³⁰⁰ No obstante, añade Rawls, la concepción política debe ayudar a que los ciudadanos, concebidos a sí mismos como libres e iguales, fortalezcan su compromiso de cooperación con otros en un sistema equitativo donde los más aventajados sean conscientes de que, con la ayuda de la distribución de las dotaciones innatas y las contingencias sociales, han logrado ocupar la posición que ostentan en la sociedad, y que por ello, tal beneficio debe redundar en la mejora de la situación de los menos aventajados.

Ahora, ¿puede el principio de utilidad restringida dar una mejor solución al descontento de los menos aventajados que la propuesta del principio de diferencia? Según

²⁹⁷ *Ibidem*, p. 170.

²⁹⁸ Al hablar de “activo común” Rawls se está refiriendo a la distribución de las dotaciones innatas y no a las dotaciones que nos pertenecen *per se* (*Ibidem*, p. 111). Además, la idea de la distribución de las capacidades innatas, en tanto acervo común, no parte de una doctrina comprensiva, sino de la idea de un pluralismo razonable. Por esta razón no se debe problematizar el “activo común” rawlsiano desde la discusión acerca de la propiedad de nuestras dotaciones innatas, ya que no es la idea que intenta proponer Rawls. Además, “son las propias personas las que poseen sus dotaciones: la integridad psicológica y física de las personas ya está garantizada por los derechos y libertades básicos que caen bajo el primer principio de justicia” (*Ibidem*, p. 112).

²⁹⁹ *Ibidem*, p. 171.

³⁰⁰ *Idem*.

Rawls, el principio de utilidad restringida le pediría a los menos aventajados recibir menores ventajas económicas y sociales en aras de las mayores ventajas para los más aventajados.³⁰¹ Esto generaría mayores tensiones de estabilidad y, por lo tanto, el ideal de reciprocidad se vería frustrado. Rawls, criticando a Harsanyi, lo expresa de la siguiente manera:

[...] en su calidad de principio de reciprocidad, el principio de diferencia se basa en nuestra disposición a devolver con la misma moneda lo que otros hacen por nosotros (o lo que nos hacen), mientras que el principio de utilidad hace más hincapié en lo que es una disposición considerablemente más débil, la de la simpatía o, mejor dicho, nuestra capacidad para identificarnos con los intereses y preocupaciones de los demás.³⁰²

Aunado a esto, la idea de un mínimo social, implícita en el principio de utilidad restringida, muestra otra debilidad frente al principio de diferencia. En *Teoría de la justicia* Rawls concibe la idea del “mínimo social” como una búsqueda del equilibrio más apropiado entre maximizar la utilidad media y asegurar un mínimo adecuado,³⁰³ sin embargo, también afirma que los que optan por este mínimo social, en realidad, se están inspirando en el principio de diferencia, lo cual lo lleva a concluir que el principio de utilidad restringida, en última instancia, no ofrecería una alternativa genuina al principio de diferencia.³⁰⁴

No obstante, a partir de algunas críticas de filósofos como Jeremy Waldron, en su ensayo “John Rawls and the Social Minimum”,³⁰⁵ Rawls reconsidera el concepto de “mínimo social” retomando la propuesta de Waldron, que lo define como “el mínimo que satisface las necesidades humanas básicas esenciales para una vida decente”,³⁰⁶ idea que Waldron relaciona con las tensiones del compromiso. Estas tensiones corresponden al hecho de que las partes se pregunten “si aquellos a los que representan son personas de las que se puede razonablemente esperar que honrarán los principios acordados de la manera requerida por la idea de acuerdo”.³⁰⁷ Por referencia a lo cual, Rawls se plantea la siguiente cuestión: “con los principios de las libertades iguales y de la equitativa igualdad de oportunidades ya adoptados,

³⁰¹ *Ibidem*, p. 173.

³⁰² *Idem*.

³⁰³ *Cfr.*, *Ibidem*, p. 174.

³⁰⁴ *Cfr.*, *Teoría de la justicia*, *Op. cit.*, pp. 291-299.

³⁰⁵ Jeremy Waldron, “John Rawls and the Social Minimum”, *Journal of Applied Philosophy*, Vol.3, No. 1 (1986), pp. 21-33. Tomado de <https://www.jstor.org/stable/24353470> (16 de junio de 2020).

³⁰⁶ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 174.

³⁰⁷ *Ibidem*, p. 145.

¿cuál es el mínimo más bajo necesario para asegurar que las tensiones del compromiso no sean excesivas?”.³⁰⁸

Las tensiones del compromiso son excesivas cuando los ciudadanos, libres e iguales, dejan de apoyar el acuerdo sustentado en los principios de justicia. De este hecho pueden seguirse dos tipos de reacción. Por un lado, existen los levantamientos, las protestas violentas, entre otros actos que muestran un descontento y resentimiento ante los principios de justicia. En este caso, los menos aventajados son los protagonistas de estas discrepancias. Otro tipo o modalidad de demostrar tal desacuerdo con los principios de justicia se expresa a través de la apatía o del sentirse rechazado por el tipo de sociedad que se ha conformado; lo que lleva, inevitablemente, a restarle importancia a los acuerdos a los que se habían comprometido bajo la concepción política de la justicia.

En una sociedad dirigida por el principio de utilidad restringida, con un mínimo social apropiado, se presentarían, en un alto grado, las tensiones del compromiso señaladas. Esto se debe a que el mínimo social –desde el principio de utilidad– es impreciso, ya que “sugiere unas directrices que no especifican un mínimo muy definido, mínimo que, en cualquier caso, dependerá en parte del nivel de riqueza de la sociedad”.³⁰⁹ Mientras que el principio de diferencia conlleva un mínimo que ayuda a maximizar, a lo largo del tiempo, las perspectivas de la vida de los menos aventajados.³¹⁰

Ahora bien, en el caso de que existiese un mínimo social definido que ayudase a cubrir todas las necesidades esenciales para una vida decente, ¿esto garantizaría que las tensiones del compromiso no fuesen excesivas? Rawls plantea que, aunque tal mínimo social no generara el primer tipo de reacción (mencionado más arriba) ante las tensiones, sí podría producirse la situación de que los menos aventajados se sintieran excluidos de la sociedad política y, por lo tanto, restaran importancia a cualquier tipo de vinculación a la misma. El profesor de Harvard sostiene que el concepto de mínimo social sería apto para un Estado capitalista de bienestar,³¹¹ en el cual, aunque se prevenga el primer modo de reacción ante

³⁰⁸ *Ibidem*, p. 174.

³⁰⁹ *Ibidem*, p. 175.

³¹⁰ *Cfr.*, *Idem*.

³¹¹ Rawls define a un “Estado capitalista de bienestar” como aquél que “rechaza el valor equitativo de las libertades políticas y, aunque muestra algún interés por la igualdad de oportunidades, no sigue las políticas necesarias para alcanzarla. Permite muy amplias desigualdades en la posesión de la propiedad real (bienes productivos y recursos naturales) de tal modo que el control de la economía y gran parte de la vida política cae

las tensiones del compromiso, los menos aventajados se sentirían excluidos y extraños a este tipo de sociedad política. Diríamos, entonces, que el principio de utilidad restringida sería entendido y aceptado en la lógica de un Estado capitalista de bienestar.

A diferencia del principio de utilidad restringida, el principio de diferencia especifica un mínimo social que se deriva de la idea de reciprocidad, antes expuesta;³¹² es decir, una sociedad donde los ciudadanos, libres e iguales, conciben el sistema social en términos de cooperación a lo largo del tiempo; así como entienden que la justicia distributiva “regula las desigualdades económicas y sociales en las perspectivas de vida, desigualdades afectadas por la clase social de origen de cada cual, por sus dotaciones innatas y por la buena fortuna a lo largo de la vida”.³¹³ En efecto, para Rawls, el mínimo social debe derivarse de una idea de reciprocidad adecuada a la sociedad política así concebida.³¹⁴

Hasta aquí hemos abordado la segunda comparación fundamental que hace Rawls entre los principios de justicia, tomados como unidad, y el principio de utilidad restringida. Aunque este último es una concepción mixta muy cercana a los principios de la justicia como imparcialidad, el filósofo estadounidense encuentra una falencia en la idea de un “mínimo social” implícito en una sociedad dirigida por el principio de utilidad restringida. Como se dijo antes, el principal problema que encuentra Rawls es la indeterminación de tal mínimo, por lo cual los menos aventajados de la sociedad podrían ser llevados a sentirse ajenos a la sociedad política debido a las grandes brechas de desigualdad que se presentarían entre los bienes sociales primarios de ingresos y de riqueza.

Por el contrario, el principio de diferencia establece un mínimo social desde la idea de reciprocidad, con la cual se aspira a que las brechas de desigualdad sean cada vez menores y a que los ciudadanos, libres e iguales, se sientan inmersos en un sistema equitativo de cooperación; en el que, a partir de la distribución de las capacidades innatas, sean conscientes

en pocas manos. Y aunque, como sugiere el nombre de ‘capitalismo del Estado de bienestar’, las provisiones de bienestar pueden ser bastante generosas y garantizar un mínimo social decente que cubra las necesidades básicas, no se reconoce ningún principio de reciprocidad que regule las desigualdades económicas y sociales” (*Ibidem*, p. 188). Para Rawls, tanto la democracia de propietarios como un régimen socialista liberal son regímenes donde los principios de la justicia como equidad ocuparían un papel central, ya que estos regímenes “garantizan las libertades básicas con el valor equitativo de las libertades políticas y la igualdad equitativa de oportunidades, y regulan las desigualdades económicas y sociales mediante un principio de mutualidad, cuando no mediante el principio de diferencia” (*Idem.*).

³¹² *Cfr.*, *Ibidem*, p. 176.

³¹³ *Ibidem*, pp. 176-177.

³¹⁴ *Cfr.*, *Ibidem*, p. 177.

del deber de ayudar a los menos aventajados. Todo lo cual conduce a concluir que, según Rawls, no sería viable que los representantes en la PO eligieran el principio de utilidad restringida por encima del principio de diferencia. Veamos ahora las críticas y los contraargumentos que presenta Harsanyi ante algunas de las principales ideas de Rawls en contra de la posibilidad de una PO utilitarista. Después trataremos de problematizar estas críticas mutuas para, de esta manera, sacar nuestras conclusiones sobre la pregunta por una sociedad equitativamente justa bajo un modelo de posición original utilitarista.

3. Crítica y contraargumentación de Harsanyi a la PO de Rawls

Para Harsanyi, uno de los errores de Rawls fue decantarse por la regla maximin como fundamento de su teoría de la justicia. En su artículo “Can the Maximin Principle Serve as a Basis for Morality?”, Harsanyi expone algunas de las paradojas que encuentra en el planteamiento de la regla maximin. Para ello, describe una situación en la que un individuo, residente de la ciudad de New York, posee dos alternativas de trabajo. En la primera alternativa el individuo tendrá que trabajar en esta ciudad, pero el trabajo será tedioso y no muy bien remunerado; en la segunda alternativa el individuo deberá viajar a Chicago y recibirá un muy buen sueldo. El hecho de optar por la segunda alternativa le implicaría al trabajador tomar un avión a Chicago, por lo que podría existir la posibilidad, aunque mínima, de que ocurriese un accidente. No obstante, al elegir la primera alternativa de empleo, el individuo tendría posibilidades nulas de un accidente aéreo. El peor resultado posible será mejor en el caso de decidir quedarse en New York, por encima de arriesgarse a morir en un accidente aéreo camino a Chicago.³¹⁵

Por consiguiente, según la regla maximin se debería elegir el trabajo en New York. Para Harsanyi, esta es una determinación irracional; si no existe una alta probabilidad de que ocurra un accidente aéreo de New York a Chicago y se tiene una fuerte preferencia por la oferta de trabajo en Chicago, el individuo no debería dudar en aceptar tal opción. El principio bayesiano conminaría al sujeto a preferir la opción con la que aseguraría la maximización de

³¹⁵ *Cfr.*, John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *The American Political Science Review*, Vol. 69, No. 2 (Jun., 1975), pp. 594-606, American Political Science Association, tomado de <http://www.jstor.org/stable/1959090>, p. 595. Todas las traducciones son mías.

su utilidad esperada. En esta lógica, actuar bajo el criterio maximin nos llevaría a un escenario donde la mejor posibilidad será quedarnos encerrados en nuestras casas fuertemente protegidos; o, incluso –como afirma Harsanyi- terminar encerrados en un manicomio,³¹⁶ ya que “es extremadamente irracional hacer que la conducta dependa totalmente de algunas contingencias altamente desfavorables a pesar de la mínima probabilidad que el individuo esté dispuesto a asignarles”.³¹⁷

Aunque, según Rawls, el principio maximin llevaría a tomar decisiones razonables en algunas situaciones, sólo serían razonables, arguye el economista húngaro, aquellas decisiones donde el criterio maximin sea equivalente a la regla de maximización de la utilidad esperada.³¹⁸ Esto se daría, por ejemplo, en el caso de plantear políticas públicas inspiradas en los dos principios descritos; para Harsanyi, las políticas dirigidas por el principio bayesiano serán vistas como las más razonables. Más adelante problematizaremos esta postura. Por ahora, sigamos analizando las principales críticas que hace el Nobel de economía al proceder rawlsiano desde la regla maximin formulada en la posición original.

Evidentemente, para Harsanyi, la regla maximin sería totalmente inadecuada para ser usada en la posición original. Los representantes ideales, ubicados en tal escenario, actuarían irracionalmente y, por lo tanto, elegirían irrazonablemente³¹⁹ bajo ignorancia e incertidumbre.³²⁰ El economista húngaro se pregunta, entonces: ¿obtendríamos un concepto de justicia satisfactorio si imagináramos los principios de justicia elegidos en la posición

³¹⁶ *Cfr., Idem.*

³¹⁷ *Idem.*

³¹⁸ *Cfr., Idem.*

³¹⁹ Para Harsanyi, siguiendo la interpretación rawlsiana, lo “racional” es lo que se espera que pudiese ser lógico desde cierta posición. Lo “razonable”, por su parte, es lo que se está dispuesto a proponer, reconocer y aceptar a partir de una decisión tomada, dentro -en el caso de Harsanyi- de la configuración del modelo de equiprobabilidad. En este caso, para Harsanyi los representantes ideales actuarían de una manera irracional porque, para él, no sería lógico usar la regla maximin para acordar unos principios realmente justos. Y, por lo tanto, la decisión que tomarían los representantes ideales en la PO rawlsiana sería irrazonable, ya que tales acuerdos estarían menoscabando los intereses de los individuos que poseen los medios necesarios para ayudar a la maximización de la satisfacción individual y, así, del bienestar social; esto a costa de darle prioridad a los menos aventajados. Es decir, en la PO de Harsanyi una gran parte de individuos no reconocerían ni aceptarían un tipo de acuerdo como al que llega Rawls en su PO, es por ello que resulta irrazonable.

³²⁰ *Cfr., Ibidem*, p. 596.

original bajo la regla maximin?³²¹ Para responder a esta cuestión, se centra en el segundo principio de justicia de Rawls, y específicamente, en el principio de diferencia.

Harsanyi arguye que el principio de diferencia conllevaría implicaciones moralmente inaceptables. Una de estas implicaciones es que tal principio nos llevaría a dar prioridad a los intereses de los individuos menos afortunados sin importar otro tipo de condiciones; incluso, bajo las más extremas condiciones.³²² Además, el filósofo utilitarista afirma que Rawls se inspira en el principio kantiano de tratar a los otros no como medios, sino como fines en sí mismos. Harsanyi pone el ejemplo de una sociedad donde existe una persona con un serio problema de retraso mental y la única forma en que puede mejorar su condición es pagando un tratamiento altamente costoso, que se obtendría privando de la mejor educación a los individuos más brillantes de la sociedad. El principio de diferencia aconsejaría concederle el costoso tratamiento al enfermo mental sin importar cuántos individuos brillantes fueran privados de una educación de alta calidad y “sin importar cuán fuertemente estos individuos desearían obtener una oportunidad como esta (y sin importar cuán grande sería la satisfacción derivada de ésta)”.³²³

Para Rawls el planteamiento utilitarista estaría usando como medios a una minoría para satisfacer los intereses de una mayoría o, más exactamente, estaría usando como medio al menos afortunado frente a los intereses y preferencias de aquéllos que podrían concederle una mayor satisfacción a la sociedad. No obstante, el economista húngaro arguye que, una sociedad, al darle prioridad al interés del individuo con el retraso mental, también estaría utilizando como medios a los individuos brillantes que no podrían tener una buena educación; y de los que, a su vez, la sociedad obtendría una mayor satisfacción.³²⁴

La intención de Harsanyi es demostrar, a partir de contraejemplos como el que hemos descrito, no sólo que la teoría de Rawls posee graves falencias en sus dos más importantes puntos de argumentación –principio maximin y el principio de diferencia–, sino que también,

³²¹ *Cfr., Idem.*

³²² *Cfr., Idem.* Aquí Harsanyi sigue la línea de algunos contraejemplos que presenta, en los que quiere dejar ver el hecho de que, por darle prioridad a los intereses de los individuos menos afortunados, se pase por alto, incluso, prioridades en materia de políticas públicas o, en último término, de bienestar social.

³²³ *Cfr., “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’ s Theory”, Op. cit., p. 597.*

³²⁴ *Cfr., Idem.*

al fracasar en estos dos ejes, es una propuesta ética errada en su totalidad. Quizá el planteamiento de la justicia como imparcialidad de Rawls, afirma el Nobel de economía, tenga implicaciones razonablemente prácticas; sin embargo, tal posibilidad también la tiene el planteamiento utilitarista al momento de confrontarlo con conclusiones razonablemente prácticas. Por ende, concluye Harsanyi, aunque Rawls no afirme que su teoría es totalmente correcta en todos los detalles que presenta, sí sostiene, no obstante, “que, al menos, los principios básicos de su teoría producen resultados más satisfactorios de lo que harían los principios básicos de las teorías utilitaristas. Sin embargo [...] (los) contraejemplos muestran, más bien, que, realmente, se da lo opuesto”.³²⁵

Ahora bien, Harsanyi describe la posición original partiendo de un ejemplo (como normalmente lo hace) acerca de un individuo que deberá elegir entre una sociedad capitalista o una socialista. El individuo basará su preferencia en sus intereses personales o teniendo en cuenta los intereses de otros individuos que pertenezcan a su círculo más cercano. Entonces, el individuo, conociendo sus preferencias, elegiría la alternativa capitalista, ya que para él habría más posibilidades de ser un aventajado bajo este sistema que bajo las condiciones del sistema socialista. Esto lo asume el autor como un juicio natural desde las preferencias personales de cada persona, ya que ésta conoce la posición en la que se encuentra ubicada y es consciente de la satisfacción que le produciría un sistema social determinado, en este caso, el capitalista.

Además, el economista húngaro habla del juicio de valor moral –distinto al juicio natural- como la decisión que tomaría el individuo si tuviese que elegir entre dos sistemas sociales sin tener conocimiento acerca de cuál sería su posición social en la sociedad en cualquier de ellos. La elección que haga el individuo estará garantizada por la misma probabilidad de tomar el lugar del individuo mejor posicionado en la sociedad; o, en su defecto, del segundo o el tercero, etc., hasta la posición del individuo peor posicionado (asunción equiprobabilística).³²⁶ Entonces, para elegir entre el sistema capitalista y el socialista, el individuo, desde la asunción equiprobabilística, usará como regla de decisión la maximización de la utilidad esperada; con esto, su elección asegurará el más alto nivel de

³²⁵ *Ibidem*, p. 598.

³²⁶ *Cfr.*, *Idem*.

utilidad promedio para cada individuo miembro de la sociedad. Y, además, se evaluarán los acuerdos sociales –instituciones, prácticas sociales, modelos económicos, etc.-desde el principio de utilidad promedio.

Lo normal es que una persona prefiera un acuerdo social por encima de otro calculando de qué manera le podría beneficiar o no tal acuerdo. La propuesta de Harsanyi está instando a los individuos a que hagan un “serio esfuerzo por ignorar”³²⁷ la información referente a la posición social donde se encuentran ubicados para, de esta forma, “hacer su elección como si ellos pensarán que tendrían la misma probabilidad de tomar el lugar de cualquier individuo particular en la sociedad”.³²⁸ El Nobel de economía agrega, además, que en esta PO el individuo tendrá dos tipos de preferencias: las personales y las morales. Las preferencias personales les darán un fuerte peso a los intereses tanto personales como a los relacionados con sus círculos sociales cercanos. Y por su parte, las preferencias morales “están basadas en un serio intento de darle el mismo peso a los intereses de cada miembro de la sociedad de acuerdo con el principio de utilidad promedio”.³²⁹

Respecto de la misma cuestión, Rawls afirma que “los juicios de probabilidad, si han de ser la base de una decisión moral, deben tener fundamentos objetivos, esto es, basarse en un conocimiento de hechos particulares (o creencias razonables)”.³³⁰ Además, en la PO “las partes ignoran los cálculos estimados de probabilidades que no estén apoyados en un conocimiento de los hechos particulares y que deriven del principio de razón insuficiente”.³³¹ Harsanyi responde a esto con dos argumentos: en el primero sostiene que la única alternativa al uso de las probabilidades subjetivas, desde la teoría de la decisión bayesiana, sería optar por un principio de decisión como la regla maximin, pero ésta –como ya lo había sostenido– conduciría a tomar decisiones irracionales en casos sumamente importantes de elección;³³²y

³²⁷ *Idem.* Resaltamos esta cita para no crear ambigüedad a partir de lo que ha afirmado Harsanyi. Esto nos servirá para algunos puntos que discutiremos respecto a esta propuesta ética.

³²⁸ *Cfr.*, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 598.

³²⁹ *Idem.*

³³⁰ John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 166.

³³¹ *Idem.* Recordemos que, según el “principio de razón insuficiente”, si no hay una razón, o explicación, suficiente para que algo sea, entonces no existirá... Es lógicamente equivalente al principio de razón suficiente (si algo existe, hay una razón o explicación suficiente de su ser) (Tomado de <http://www.philosophica.info/voces/razon-suficiente/Razon-suficiente.html>) (01 de junio de 2020).

³³² *Cfr.*, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 599.

en un segundo argumento, el economista húngaro afirma que, según la teoría de decisión bayesiana, “cualquier individuo que tenga que decidir, cuya conducta sea consistente con algunos, muy convincentes, postulados racionales, simplemente no puede evitar actuar como si usara probabilidades subjetivas”,³³³ es decir, “no puede evitar actuar como si él tratara de maximizar su utilidad esperada, computada sobre la base de algún grupo de probabilidades subjetivas”.³³⁴ Además, Harsanyi arguye que, incluso bajo la regla maximin, un individuo estaría abocado a no evitar elegir fuera de probabilidades subjetivas, al menos implícitamente.³³⁵ Es decir, aunque tal individuo evite pensar en probabilidades, “su conducta equivaldrá, realmente, a asignar una probabilidad a la peor posibilidad en cualquier caso dado”.³³⁶

No obstante, para Rawls una determinada acción en la PO será más fácil de justificar si estas acciones están basadas en la regla maximin en lugar de en la asunción equiprobabilística;³³⁷ Harsanyi responde a esta aseveración afirmando que la asunción equiprobabilística puede ser justificada por el “principio de indiferencia” (*principle of indifference*), según el cual, en la ausencia de cualquier razón para esperar un evento u otro, se debería asignar la misma probabilidad a todos los posibles eventos.³³⁸ Así como también está basada en el “principio moral de asignar, *a priori*, el mismo peso a cada interés individual”.³³⁹ Además, según Harsanyi, usar la regla maximin en la PO equivale a asignarle una probabilidad a la posibilidad de que un individuo pueda terminar ubicado en la peor posición de la sociedad, lo cual no tendría ningún tipo de justificación racional.³⁴⁰

Como lo abordamos antes, en la teoría ética utilitarista de Harsanyi las funciones de utilidad de von Neumann-Morgenstern de los individuos hacen parte de la definición de justicia y otros valores morales.³⁴¹ No obstante, para Rawls, al usar los postulados de

³³³ *Idem.*

³³⁴ *Idem.*

³³⁵ *Cfr., Idem.*

³³⁶ *Idem.*

³³⁷ *Cfr., Teoría de la justicia, Op. cit., p. 163.*

³³⁸ *Cfr.,* <https://www.collinsdictionary.com/dictionary/english/principle-of-indifference> (29 de mayo de 2020). La traducción es mía.

³³⁹ John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’ Theory, *Op. cit.*, p. 599.

³⁴⁰ *Cfr., Idem.*

³⁴¹ *Cfr., Ibidem, p. 600.*

Neumann-Morgenstern en las decisiones sociales, “los sentimientos de las personas acerca de correr riesgos afectarán el criterio de bienestar que ha de maximizarse”;³⁴² por lo que, para Rawls, sería conveniente no usar este tipo de funciones de utilidad en las definiciones de valores morales al momento de construir un concepto político de justicia. Harsanyi responde a esta crítica afirmando que la opinión de Rawls está basada en una mala interpretación de la función de utilidad de von Neumann-Morgenstern; es decir, aunque esta función es estimada desde la conducta de elección bajo riesgo e incertidumbre de un individuo, no significa, sin embargo, que tal función de utilidad sea, simplemente, un indicador de las actitudes de tal individuo al momento de correr riesgos.³⁴³ Más bien, como una función de utilidad cardinal, la principal tarea de una función de utilidad de von Neumann-Morgenstern es indicar cuánta utilidad –por ejemplo, cuánta importancia subjetiva– un individuo asigna a varias de sus metas.³⁴⁴ Por ende, “las funciones de utilidad de von Neumann-Morgenstern tienen un legítimo lugar en la ética, ya que tales funciones expresan la importancia subjetiva que las personas ponen a sus distintas necesidades e intereses”.³⁴⁵

Otra de las principales críticas de Rawls a Harsanyi se refiere a la posición de este último respecto de las comparaciones interpersonales de utilidad. Específicamente, Rawls asevera que las utilidades representan la validez de las alternativas para el individuo que está tomando la decisión –el observador imparcial y simpático–, estimadas en su tabla de valores.³⁴⁶ Sin embargo, “hay tantas personas diferentes como utilidades”.³⁴⁷ En efecto, el filósofo estadounidense se pregunta “¿cómo deberemos valorar el modo de vida y sistema de objetivos formales del otro?, ¿conforme a sus objetivos o conforme a los nuestros?”.³⁴⁸ Los conflictos de intereses, que se generan a partir de las circunstancias de la justicia, son una muestra de que el valor que tiene un determinado fin para un individuo es totalmente diferente del que tiene para otro: “las demandas conflictivas no sólo surgen porque la gente quiere tipos de cosas semejantes para satisfacer deseos semejantes [...] sino porque sus

³⁴² John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 298.

³⁴³ *Cfr.*, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 600.

³⁴⁴ *Cfr.*, *Idem.*

³⁴⁵ *Idem.*

³⁴⁶ *Cfr.*, *Teoría de la justicia*, *Op. cit.*, p. 167.

³⁴⁷ *Idem.*

³⁴⁸ *Idem.*

concepciones del bien difieren”.³⁴⁹ He aquí la imposibilidad, según Rawls, para hacer uso de las comparaciones interpersonales al momento de construir una concepción de justicia como equidad.

Harsanyi responde con el señalamiento de que el uso de las comparaciones interpersonales de utilidad en su propuesta ética atañe al hecho de que el individuo “debe tratar de estimar qué nivel de utilidad disfrutaría si él mismo estuviese ubicado en las condiciones físicas, económicas y sociales objetivas de cualquier otro individuo”.³⁵⁰ Una manera de ejemplificar esta posición de Harsanyi es al momento de juzgar si cierta política estatal es justa o no respecto a no tomar alcohol pasadas las 10 pm. Mi juicio respecto a esta política pública dependerá de si a los integrantes de mi sociedad les gusta o no tomar alcohol pasadas las diez de la noche, para lo cual sería necesario establecer comparaciones interpersonales de utilidad. Sería insensato proferir un juicio en mis propios términos; en este caso, no puedo decir que, por el hecho de que a mí me guste tomar alcohol pasadas las diez de la noche, la política pública dispuesta es injusta. Para Harsanyi, además, “las preferencias y funciones de utilidad de todos los individuos están gobernadas por las mismas leyes psicológicas”.³⁵¹ Es decir, aunque mi función de utilidad sea diferente a la de otro, ambas funciones están gobernadas por este tipo de leyes básicas.³⁵² Esto significa que:

bajo condiciones ideales, por ejemplo, si tuviésemos conocimiento total de las leyes psicológicas que gobiernan a las preferencias de los individuos y sus funciones de utilidad, y si contáramos con la suficiente información sobre las características personales de los otros individuos, entonces, podríamos hacer unas comparaciones interpersonales de utilidad completamente libres de error.³⁵³

Harsanyi afirma que, aunque Rawls exprese sus dudas y desacuerdos frente a las comparaciones interpersonales de utilidad, no obstante, no intenta refutar la teoría que el economista húngaro propone acerca de tales comparaciones. Antes bien, se limita a proponer una alternativa a las comparaciones interpersonales –bienes sociales primarios– que, en

³⁴⁹ *Idem.*

³⁵⁰ John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 600.

³⁵¹ *Idem.*

³⁵² *Cfr., Idem.*

³⁵³ *Ibidem*, p. 601.

opinión de Harsanyi, nos llevaría a caer en postulados morales altamente arbitrarios e improbables.³⁵⁴

Una última observación que hace Harsanyi respecto a la postura de Rawls frente al utilitarismo concierne a la estabilidad social. Para Rawls, sería muy difícil que una sociedad gobernada por un principio de utilidad –como en el caso de la concepción mixta desde el principio de utilidad restringida– sea estable. Especialmente, esto se vería reflejado en la ida de reciprocidad que refleja el principio de diferencia y de la cual adolece el principio de utilidad. Como habíamos mencionado, los menos aventajados serían los principales autores de tal inestabilidad en un Estado dirigido por las lógicas del principio utilitarista. No obstante, Harsanyi arguye que esta presunción de estabilidad es ingenua, ya que el tipo de sociedad que describe Rawls nunca ha existido y es totalmente desconocida por cualquier estudioso del tema. Aunque hemos vivido, como sociedad, momentos de acercamiento a este tipo de ideales, es verídico, desde la postura del economista húngaro, que vivimos en sociedades donde los intereses egoístas y altruistas conviven. Por lo tanto, para Harsanyi, “cualquier sociedad saludable necesita un apropiado balance entre motivaciones egoístas y altruistas”.³⁵⁵ Es decir, una sociedad justa, con un nivel razonable de estabilidad social, no sería una sociedad donde todos los ciudadanos estarían motivados por un sentido de la justicia, sino, más bien, aquélla donde la mayoría de las personas estarían motivadas por la mixtura normal entre intereses egoístas y altruistas.³⁵⁶

Hasta aquí, en este capítulo hemos rastreado las distintas alteraciones que sufriría la PO dirigida bajo un principio utilitarista. Para ello, expusimos la manera en que se encuentra configurada la PO desde la concepción de la justicia como equidad. Resaltamos, especialmente, las constricciones formales y el velo de la ignorancia, elementos que ayudan a establecer criterios bajo los cuales se sostiene la idea de una sociedad imparcialidad justa. De igual manera, expusimos los principales argumentos que propone Rawls para elegir los dos principios de justicia por encima del principio de utilidad media. Para ello, recurrimos a las dos comparaciones fundamentales que establece el profesor de Harvard en *La justicia como equidad*, lo cual nos condujo a ubicar los puntos clave en los que diverge el principio

³⁵⁴ Cfr., *Idem*.

³⁵⁵ *Ibidem*, p. 604.

³⁵⁶ Cfr., *Idem*.

de utilidad media de la justicia como equidad. Establecimos, por ende, el hecho de que una sociedad dirigida por un principio de utilidad adolecería tanto de la idea de igualdad como de la idea de reciprocidad.

Asimismo, después de presentar el edificio argumentativo rawlsiano desde la PO y las principales críticas al principio de utilidad media, abordamos las críticas y contraargumentos que elabora Harsanyi en respuesta a Rawls, resaltando las razones por las que ideas como la regla maximin y el principio de diferencia serían rechazadas en una PO utilitarista. De igual manera, retomamos críticas puntuales de Rawls a la propuesta ética de Harsanyi para poner de relieve la manera en que el economista húngaro puede responder a estas críticas desde la lógica de su propuesta utilitarista.

En este momento, con los suficientes recursos argumentativos, procedentes tanto de Rawls como de Harsanyi, podemos establecer nuestras conclusiones respecto a la posibilidad de una sociedad equitativamente justa bajo un modelo de PO utilitarista.

Conclusiones

Luego de haber expuesto los principales argumentos del utilitarismo, tanto en su doctrina clásica como contemporánea, contamos con herramientas suficientes para dirimir si es posible hablar de una sociedad justa bajo un modelo de *posición original utilitarista*. En función de ello, además, retomamos la teoría de la justicia de John Rawls centrándonos en la manera en que éste se distancia de los postulados centrales de la doctrina utilitarista para pensar las mejores condiciones posibles que lleven a la construcción de una sociedad imparcialmente justa. Destacamos, específicamente, la confrontación de las ideas del economista húngaro John Harsanyi con la argumentación rawlsiana acerca de la imposibilidad de una sociedad justa dirigida por una PO utilitarista.

A lo largo de nuestra argumentación hemos constatado que la configuración de una PO utilitarista es insuficiente para conformar una sociedad justa. Esta conclusión resulta del contraste realizado entre la PO utilitarista, desde los argumentos de Harsanyi a partir del modelo de equiprobabilidad de los juicios de valor moral, y la PO rawlsiana, centrada en las condiciones bajo las cuales los representantes ideales eligen guiados por la regla maximin bajo criterios de igualdad y reciprocidad. Este análisis lo llevamos a cabo desde dos

comparaciones fundamentales propuestas por Rawls, en las que se prueba por qué sería más razonable, para las partes en la PO, elegir los dos principios de justicia por encima del principio de utilidad.

El concepto de sociedad justa del cual hemos partido en este trabajo se fundamenta en la propuesta rawlsiana de una sociedad equitativamente cooperativa, donde resaltan, como ejes centrales, tanto la igualdad como la reciprocidad. Estos ejes reflejan, por un lado, tanto los derechos y las libertades básicas como las oportunidades igualmente equitativas; y, por otro lado, en el caso de la idea de reciprocidad, el principio de diferencia. En este sentido, no sería viable fundar una sociedad –que pretende ser imparcialmente justa- bajo una PO utilitarista, ya que, en la búsqueda de la maximización de la suma del bienestar social, se verían sacrificadas tanto la igualdad como la reciprocidad.

La regla maximin, como directriz en la elección de los principios de justicia rawlsianos, es parte de la primera comparación que establece Rawls para demostrar que, si tomáramos como guía de elección el modelo de equiprobabilidad de Harsanyi, sería altamente posible que en la sociedad que surgiera de tal escenario se permitieran todo tipo de injusticias en contra de los ciudadanos menos aventajados. Es por ello que Rawls, descartando un razonamiento desde la probabilidad, pone a las partes en la PO a elegir desde un nivel garantizable, esto es, a evaluar las distintas alternativas desde sus peores resultados posibles y a optar por aquella cuyo peor resultado sea el mejor.

Los representantes en la PO rawlsiana, bajo el velo de la ignorancia, se encuentran en una situación de incertidumbre más que de riesgo, ya que no existe una base objetiva por la cual poder estimar las distintas probabilidades al llevar a cabo la elección de los principios de justicia. En la PO utilitarista, desde Harsanyi, se tiene como referencia la teoría de la elección racional en circunstancias de riesgo e incertidumbre. Desde este marco, para Harsanyi resulta más razonable, por lo tanto, que el observador imparcial y empático asuma que tendrá la misma posibilidad de ocupar cualquier posición social en el tipo de sociedad que elija en el marco de una PO utilitarista; esto implicará, además, que el observador imparcial prefiera aquella opción que asegurará la maximización de su utilidad esperada. En efecto, para el Nobel de economía, sería irracional pretender elegir principios de justicia en

la PO desde la regla maximin, pues sería irrazonable que las principales directrices de una sociedad estuviesen guiadas por un miedo irracional a la incertidumbre.

Rawls considera que dentro del razonamiento que establece en la regla maximin no se descarta la maximización de la utilidad esperada de los individuos; no obstante, esta debe estar al servicio de un ordenamiento de los intereses fundamentales de los ciudadanos en condiciones de igualdad, y es en este punto donde el asegurar un nivel garantizable resulta necesario para establecer los mejores términos de justicia. Por lo tanto, para Rawls es imprescindible que una concepción política de la justicia garantice la defensa de los derechos y libertades básicos, oportunidades equitativas y una adecuada distribución del ingreso y la riqueza. En esta lógica, los dos principios de justicia rawlsianos serían los más favorables para hablar de una sociedad imparcialmente justa, por encima del establecimiento de un principio de utilidad.

En una segunda comparación fundamental, Rawls confronta los dos principios de justicia con el principio de utilidad restringida. Este último resulta de reemplazar, en el marco de los principios de justicia rawlsianos, el principio de diferencia por el principio de utilidad media. La dificultad que encuentra Rawls en este principio mixto de justicia es que la existencia de un mínimo social apropiado (que supondría la aplicación del principio de utilidad restringida a la estructura básica de la sociedad) llevaría a serios inconvenientes en la estabilidad de una sociedad que pretende ser imparcialmente cooperativa. En la primera comparación se pudo demostrar por qué la igualdad resultaría gravemente afectada desde la lógica del principio de utilidad; en esta segunda comparación se hace énfasis, más bien, en la importancia de la reciprocidad, y, por lo tanto, en la amenaza que podría representar para una sociedad orientarse por el principio de utilidad restringida.

Ante la existencia tanto de los menos aventajados como de los más aventajados en una sociedad democrática, Rawls considera importante un principio que ayude a regular las grandes desigualdades de ingresos y riqueza entre unos y otros. La respuesta la encuentra en el principio de diferencia, que expresa y defiende la idea de reciprocidad. Esta idea también genera un ambiente de *estabilidad* de los individuos respecto de los términos de cooperación social acordados, ya que el principio de diferencia garantiza que los menos aventajados no quedarán insertos en un margen abismal en medio de la distribución de los ingresos y la

riqueza en la sociedad. Lo mismo sucede con los más aventajados, para los cuales el hecho de cooperar les significará beneficios sin necesidad de olvidarse de aquéllos que no tuvieron la misma suerte que ellos.

El principio de utilidad restringida, a juicio de Rawls, conllevaría significativas tensiones en la estabilidad social. Esta fractura en la estabilidad social se debería a que, por un lado, el principio de utilidad restringida le concedería menores ventajas económicas a los menos aventajados a costa del mayor beneficio de los más aventajados. Por otro lado, el mínimo social, que se pretende asegurar desde tal principio, es impreciso y siempre quedaría sujeto a la acumulación indiscriminada de la riqueza. El principio de diferencia, por su parte, consta de un mínimo que actúa en función de la maximización, a lo largo del tiempo, de las perspectivas de los menos aventajados de la sociedad. Es más, aunque el principio de utilidad restringida llegara a establecer un mínimo social definido, eso no sería garantía para asegurar una estabilidad social y, por lo tanto, la ausencia de tensión en los compromisos acordados en una sociedad democrática. La principal distinción se ubica, entonces, en el hecho de que el principio de diferencia se inspira en la reciprocidad, mientras que el principio de utilidad restringida pretende asegurar la mayor acumulación de riquezas posibles, sin importar que a los menos aventajados les corresponda cada vez menos.

Para Harsanyi, orientarse por el principio de diferencia significaría una espada de doble filo para una sociedad que pretende ser imparcialmente justa. Arguye que, por priorizar a los menos aventajados, se terminarían usando como medios a aquellos individuos que podrían significarle un mayor beneficio a la sociedad. Por ende, para Harsanyi será importante que el observador imparcial elija desde un juicio de valor moral –considerando la existencia de sus preferencias personales y morales- que tenga en cuenta la asunción probabilística antes mencionada.

Otro de los puntos que objeta Rawls a la propuesta ética de Harsanyi es el referente a las comparaciones interpersonales de utilidad. No podríamos pretender, según la argumentación rawlsiana, calcular –desde nuestros propios criterios de utilidad- el modo de vida y el sistema de objetivos formales del otro, ya que cada individuo es diferente. Construir una concepción política de justicia a partir de las comparaciones interpersonales de utilidad podría conllevar escenarios en los que se desconocerían intereses fundamentales que una

sociedad justa debería garantizar a todo individuo. Para Harsanyi, la crítica que le dirige Rawls a propósito de las comparaciones interpersonales de utilidad no corresponde a lo que quiso establecer en su teoría utilitarista; esto es, Rawls se basa en un concepto de comparaciones interpersonales de utilidad clásico, dejando de lado la reformulación que lleva a cabo Harsanyi de estas comparaciones a partir de su propuesta de equiprobabilidad. En efecto, el economista húngaro sostiene y defiende que tales comparaciones sirven para que un individuo trate de imaginar qué nivel de utilidad obtendría estando ubicado bajo las circunstancias determinadas y objetivas de otro individuo, de acuerdo con los “postulados de similitudes”; lo cual, según Harsanyi, conduciría a que las políticas públicas adoptadas sean justas.³⁵⁷

Rawls también critica el uso que hace Harsanyi de la función de utilidad de von Neumann-Morgenstern dentro de su modelo de equiprobabilidad, ya que el criterio de bienestar que se buscaría maximizar resultaría afectado por la fuerte influencia de la toma de riesgos en los individuos al momento de elegir. Para Harsanyi, el análisis de Rawls es sesgado, ya que la función de utilidad de von Neumann-Morgenstern no sólo funciona como indicador de la actitud del individuo frente al riesgo, sino que su principal papel consiste en mostrar la importancia que cada individuo le concede a sus necesidades e intereses particulares, lo cual no representaría problema alguno al momento de construir un concepto político de justicia.

Como hemos corroborado, una PO utilitarista no tendría lugar en los términos en que Rawls describe su PO, con base en la idea de una sociedad imparcialmente justa. Esto resulta evidente en la disputa irreconciliable de argumentos que hemos expuesto. Consideramos, además, que tal disputa es el resultado del modo en que cada autor sueña con un tipo de sociedad específico. De esta manera, como lo describe Rawls en *Teoría de la justicia*, existe una diferencia esencial entre su teoría y el utilitarismo. El profesor de Harvard se refiere al hecho de que “el utilitarismo es una teoría teleológica, mientras que la justicia como

³⁵⁷ En este punto es importante tener presente el concepto de comparaciones interpersonales de utilidad que propone Sidgwick, así como el concepto de comparaciones interpersonales de utilidad que establece Harsanyi, el cual se ha explicado con detenimiento en el apartado “comparaciones interpersonales de utilidad y el postulado de similitudes en Harsanyi”. De esta forma podremos entender mejor por qué Harsanyi sostiene que Rawls le hace la crítica desde un trasfondo teórico utilitarista clásico.

imparcialidad no lo es”,³⁵⁸ ésta sería una teoría deontológica. Es decir, en el utilitarismo “la satisfacción de cualquier deseo tiene algún valor en sí, valor que deberá tomarse en cuenta al decidir lo que es justo”³⁵⁹; mientras que en la justicia como imparcialidad “las personas en la posición original escogerían el principio de igual libertad y restringirían las desigualdades económicas y sociales a las que fueran en interés de todos”³⁶⁰, por lo que no habría “razón para pensar que las instituciones justas maximizarán el bien”.³⁶¹

Sumado a lo anterior, para Michael Moehler, en su artículo “The Rawls-Harsanyi dispute: a moral point of view”,³⁶² la disputa entre el utilitarismo de Harsanyi y el concepto político de justicia de Rawls no se refiere tanto a un problema de aplicación de una teoría de la decisión normativa a la PO rawlsiana, sino, más bien, a que cada autor presupone ideas morales concernientes a la justicia.³⁶³ Por su parte, Will Kymlicka, en su artículo “Rawls on Teleology and Deontology”,³⁶⁴ afirma que la discusión entre el utilitarismo y Rawls se centra en el cómo se podría llegar a dar igual peso a cada demanda o interés personal (*personal claims*) a partir de la formulación de unos principios de justicia.³⁶⁵ El análisis de Kymlicka nos ayuda a corroborar por qué, a partir de la propuesta rawlsiana, sería inviable pensar una sociedad –que pretende ser imparcialmente justa- bajo parámetros utilitaristas; como lo habíamos señalado arriba, bajo una PO utilitarista los intereses fundamentales de un número significativo de individuos se verían sacrificados por la obtención del mayor bienestar en una determinada sociedad.

Ahora bien, además de los irreconciliables puntos de vista que defiende cada autor, es necesario detenerse también en la manera en que Harsanyi plantea algunas de las principales críticas a Rawls. Específicamente, en el hecho de que el economista húngaro parece tomar ciertos argumentos rawlsianos descontextualizados del edificio argumentativo que el profesor de Harvard consideró imprescindible para entender su teoría. Esto tiene como

³⁵⁸ John Rawls, *Teoría de la justicia*, *Op. cit.*, p. 40.

³⁵⁹ *Ibidem*, p. 41.

³⁶⁰ *Idem*.

³⁶¹ *Idem*.

³⁶² Moehler, Michael, “The Rawls-Harsanyi dispute: a moral point of view” in *Pacific Philosophical Quarterly*, *Volumen 99*, Issue 1, pp. 82-99, 2018, Tomado de <https://doi.org/10.1111/papq.12140> (07 de octubre de 2020)

³⁶³ *Ibidem*, p. 84.

³⁶⁴ Kymlicka, Will, “Rawls on teleology and deontology” in *Philosophy & public affaris*, *Vol. 17*, No. 3, (Summer, 1988), pp. 173-190, Tomado de <http://www.jstor.org/stable/2265243> (27 de octubre de 2020).

³⁶⁵ *Ibidem*, p. 184.

efecto que las críticas de Harsanyi se tornen débiles frente al conjunto de la propuesta rawlsiana. Profundicemos, ahora, en algunos de los matices problemáticos de la crítica del economista húngaro a la regla *maximin* y al principio de diferencia.

Uno de los argumentos empleados por Harsanyi para demostrar por qué la regla *maximin* no sería apropiada para tomarla como regla de decisión en la PO, es que ésta llevaría a que el aumento en la satisfacción o el bienestar de un individuo sea sacrificado en beneficio de la prioridad concedida al individuo menos afortunado. El economista húngaro presenta un contraejemplo donde, en una sociedad conformada por dos individuos, un individuo A tiene una excepcional capacidad para las matemáticas y desea, fuertemente, estudiar una especialización en matemáticas avanzadas. Mientras que un individuo B sufre de un severo retardo mental y necesita un tratamiento de recuperación. La sociedad cuenta con un excedente de recursos que pueden ser destinados sólo a una de las necesidades de estos dos individuos, en el supuesto de que no puede dividirlos. Se debe tener en cuenta que el tratamiento de recuperación que se le aplicaría al individuo B se traduciría en pequeñas mejorías; en efecto, sería un avance, pero le traería poca satisfacción a éste. En cambio, la satisfacción del individuo A, luego de especializarse en matemáticas avanzadas, sería mucho mayor que la obtenida por el individuo B. Para Harsanyi el principio de diferencia, desde la regla *maximin*, llevaría a que el excedente con el que cuenta la sociedad fuera destinado al individuo B. Por el contrario, desde la teoría utilitarista se le facilitarían los recursos al individuo A para que lleve a cabo sus estudios de alto nivel y, por ende, obtenga una mayor satisfacción y bienestar en esa determinada sociedad.³⁶⁶

A este respecto, una premisa que establece Rawls para explicar la regla *maximin* es que “el argumento guiado por la regla se ajusta a la idea de que los agentes racionales maximizan su utilidad esperada, pero solo si esa utilidad esperada se concibe vacía de contenido sustantivo”.³⁶⁷ Como se observa, uno de los principales rasgos del contraejemplo que acabamos de mostrar es que Harsanyi está evaluando la regla *maximin* desde un contenido sustantivo específico, en este caso, la satisfacción o el bienestar social. Por ende, resulta inconsistente la crítica del economista húngaro, ya que Rawls no hace depender la

³⁶⁶ *Cfr.*, John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 596.

³⁶⁷ John Rawls, *La justicia como equidad*, *Op. cit.*, p. 140.

eficacia del uso de la regla *maximin* en la PO de la obtención de la mayor satisfacción o placer esperado. Antes bien, como habíamos apuntado antes, la utilidad esperada, como idea puramente formal, “representa simplemente el orden, o clasificación, en que las alternativas se juzgan mejores o peores en punto a satisfacer los intereses fundamentales de los agentes, que en este caso son los intereses de ciudadanos libres e iguales”.³⁶⁸

Además, Harsanyi está proponiendo asignar, en el caso mencionado, una cantidad de dinero que la sociedad tendría como excedente al individuo A o al individuo B. Y en otros ejemplos de Harsanyi, también los bienes específicos a distribuir son dinero, medicina, etc..³⁶⁹ Sin embargo, evaluar críticamente la regla *maximin* y el principio de diferencia desde la argumentación rawlsiana según este tipo de distribución resulta inconsistente, ya que, para Rawls, la configuración misma de la PO lleva a usar la regla *maximin* dentro del marco de lo que se propondrá en el principio de diferencia desde la premisa de que:

El problema de la justicia social no consiste en distribuir “*ad libitum*” entre ciertas personas diversas cantidades de algo, sea dinero, propiedades o cualquier otra cosa [...] [Antes bien] los dos principios están vinculados formando una concepción de la justicia que se aplica a la estructura básica de la sociedad en su conjunto.³⁷⁰

En el artículo al que estoy refiriéndome, Harsanyi afirma que cuando los contraejemplos están dirigidos precisamente contra los principios más fundamentales de la teoría, como la regla *maximin* y el principio de diferencia en el caso de la teoría de Rawls, “si los contraejemplos son válidos, eso solo significa que la teoría está fundamentalmente errada”.³⁷¹ No obstante, como se ha demostrado, los contraejemplos que presenta Harsanyi pretendiendo criticar tanto la regla *maximin* como el principio de diferencia, se hallan descontextualizados y, por ende, no demuestran consistente de una posible falla en la teoría rawlsiana.

Ahora bien, a pesar de que las respuestas de Rawls a las críticas de Harsanyi parecen ser sólidas consideramos que existen algunas inconsistencias en la segunda comparación que hace el filósofo estadounidense. Específicamente, parece que Rawls se queda corto en cuanto

³⁶⁸ *Ibidem*, pp. 140-141.

³⁶⁹ *Cfr.*, John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 596. En esta parte del artículo de Harsanyi pone un contraejemplo donde evalúa la justicia en una determinada sociedad a partir de cómo hace la distribución de un determinado antibiótico entre dos individuos con diferentes rasgos y necesidades.

³⁷⁰ John Rawls, *Teoría de la justicia*, *Op. cit.*, pp. 153-154.

³⁷¹ *Cfr.*, John Harsanyi, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *Op. cit.*, p. 597.

a criticar el “mínimo social apropiado” que, según él, un principio utilitarista propondría en reemplazo del principio de diferencia. Rawls afirma que una debilidad del principio de utilidad restringida es la inexistencia de un mínimo social definido; lo cual conducirá, inevitablemente, a las tensiones en la estabilidad social. No obstante, aunque el principio de diferencia establezca un nivel garantizable desde lo propuesto en la regla maximin, también es muy cierto que el mismo Rawls admite que existirá una indeterminación en las porciones distributivas que se vayan creando, a lo largo del tiempo, a partir de la cooperación productiva de los miembros de la sociedad³⁷²; es decir, podría ser altamente probable que, en su aplicación, el principio de diferencia termine convirtiéndose en una versión muy cercana al mínimo social del cual se habla en el principio de utilidad restringida.

Por consiguiente, es evidente que la teoría rawlsiana presenta una seria falencia en lo que atañe al planteamiento del principio de diferencia y a su aplicación en nuestras actuales democracias. Muchos de los países democráticos actuales se acercan mucho a la descripción de Rawls de un Estado capitalista de bienestar. La propuesta utilitarista es muy llamativa ante los deseos de pocos individuos de acumular, cada vez más, ingentes riquezas sin considerar —o incluso usando como carne de cañón— las expectativas sociales de los menos aventajados. Al presentarse, entonces, el principio de diferencia como una solución razonable frente a las implicaciones que supondría la aplicación del criterio utilitarista, parece también necesario revisar tal principio buscando encontrar la mejor forma de que no quede convertido, al final, en una variante del principio utilitarista.

Es racional que todos busquemos la máxima satisfacción de nuestro propio bien y también es racional que trabajemos en la búsqueda del mayor bienestar de todos en la sociedad. Ningún individuo racional estaría en contra de estas dos premisas. Sin embargo, también es cierto que la manera en que nos hemos ido constituyendo como sociedades democráticas a lo largo del tiempo nos ha mostrado que no basta con poseer este tipo de deseos racionales cuando se trata de construir una democracia justa y equitativa. Por ello es necesario buscar el camino más razonable para que, a través de las principales instituciones políticas y sociales de un Estado social de derecho con régimen constitucional, se garanticen intereses fundamentales e inviolables, que cada individuo pueda hacer propios y respetar en

³⁷² Cfr., John Rawls, *La justicia como equidad*, *Op. cit.*, p. 103.

términos de libertad, igualdad y reciprocidad. Por esta razón la teoría de la justicia de John Rawls no ha perdido vigencia y resuena como un grito ahogado en medio de las innumerables violaciones a los derechos y libertades básicos, frente a la ausencia de oportunidades en países con grandes recursos para fomentar la cooperación productiva donde las brechas abismales entre ricos y pobres son cada vez más grandes. Sin embargo, también somos conscientes de que la propuesta rawlsiana, en su formulación –especialmente, desde el principio de diferencia- necesita ser revisada y mejorada para así acercarnos a la mejor puesta en escena posible de una sociedad imparcialmente justa.

Fuentes documentales

Araújo, Cícero, “Bentham: el utilitarismo y la filosofía política moderna” en *La filosofía política moderna. De Hobbes a Marx*, CLACSO, Consejo Latinoamericano de Ciencias Sociales, Buenos Aires, 2000, Tomado de <http://bibliotecavirtual.clacso.org.ar/clacso/se/20100609023007/11cap10.pdf>

Bidet, Jacques, *John Rawls y la teoría de la justicia*, Bellaterra, Barcelona, 2000.

Brandt, Richard, *Ethical Theory: The Problems of Normative and Critical Ethics*, Prentice-Hall, Inc, Englemood Cliffs, NJ, 1959.

Brandt, Richard, *A Theory of the Good and the Right*, Clarendon Press Oxford 1979.

Carrington Heat, Will, “von Newmann/Morgenstern Decision Making and Harsanyi’s Theory of Justice: after it Riks Personal Freedom and Individual Liberty, Can It Be Really Just?”, *The American Journal of Economics on Sociology*, Vol. 47, No. 3 (Jul., 1988), pp. 355-362. Tomado de <https://www.jstor.org/stable/3486484>

Da Silveira, Pablo, *John Rawls y la justicia distributiva*, Campo de Ideas, Madrid, 2013.

Dimmock, Mark y Fisher, Andrew, “Utilitarianism” en *Ethics for A-level*, Open Book Publishers, Tomado de <https://www.jstor.org/stable/j.ctt1wc7r6j.5>

Fleming, Marcus, “A Cardinal Concept of Welfare”, *The Quarterly Journal of Economics*, Vol. 66, No. 3 (Aug. 1952), pp. 366-384. Oxford University Press, Tomado de <https://jstor.org/stable/1885309>

- Freeman, Samuel, *Rawls*, Fondo de Cultura Económica, México, 2007.
- Gargarella, Roberto, *Las teorías de la justicia después de Rawls: un breve manual de filosofía política*, Paidós, Barcelona, 1999.
- Gauthier, David, *La moral por acuerdo*, Gedisa, Barcelona, 1986.
- Harsanyi, John, “Can the Maximum Principle Serve as a Basis for Morality? A Critique of Jhon Rawls’s Theory”, *The American Political Science Review*, Vol. 69, No. 2 (Jun., 1975), pp. 594-606, American Political Science Association, tomado de <http://www.jstor.org/stable/1959090> el 22 de febrero de 2020.
- Harsanyi, John, “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility”, *The Journal of Political Economy*, Vo. 63, No. 4 (Aug. 1955), pp. 309-321, University of Chicago Press, Accessed March 12, 2020. <http://www.jstor.org/stable/1827128>
- Harsanyi, John, “Morality and the Theory of Rational Behaviour”, *Social Research* 44, no. 4 (1977): 623-56. Accessed March 15, 2020. www.jstor.org/stable/40971169, p. 623. La traducción es mía.
- Herzog, Don, “Against utilitarianism, Justification in Political Theory”, *Cornell University Press*, 1985, Tomado de <https://www.jstor.org/stable/10.7591/j.ctt207g6t3.7>
- Kukathas, Chandran y Pettit, Philip, *La teoría de la justicia de John Rawls y sus críticos*, Tecnos, Madrid, 2004.
- Kymlicka, Will, “Rawls on teleology and deontology” en *Philosophy & public affaris*, Vol. 17, No. 3, (Summer, 1988), pp. 173-190, Tomado de <http://www.jstor.org/stable/2265243>
- Locke, John, *Ensayos sobre el entendimiento humano*, Fondo de Cultura Económica, México, 1956.
- Mill, John Stuart, *El utilitarismo*, Alianza Editorial, Madrid, 1984.
- Mill, John Stuart, *Utilitarianism*, Global Grey ebooks, 2015. Tomado de <https://www.globalgreybooks.com/utilitarianism-ebook.html>, (20 de diciembre de 2019).
- Mill, John Stuart, *Bentham*, Estudio preliminar, traducción y notas: Carlos Mellizo, Tecnos, 1993

Moehler, Michael, "The Rawls-Harsanyi dispute: a moral point of view", *Pacific Philosophical Quarterly*, Volumen 99, Issue 1, pp. 82-99, 2018, Tomado de <https://doi.org/10.1111/papq.12140> (07 de agosto de 2020)

Nozick, Robert, *Anarquía, estado y utopía*, Fondo de Cultura Económica, México, 1974.

Pogge, Thomas, "John Rawls: una biografía" en *Revista Co-herencia Vol. 7, No. 12* Enero – Junio, Medellín, 2010, pp. 13-42.

Puyol, Ángel, *Rawls: el filósofo de la justicia*, Batiscafo, Barcelona, 2015.

Rawls, John, *La justicia como equidad: una reformulación*, Paidós, Barcelona, 2000.

Rawls, John, *Teoría de la justicia*, Fondo de Cultura Económica, México, 1997.

Rawls, John, *Liberalismo político*, Fondo de Cultura Económica, México, 1995.

Rawls, John, "Justice as fairness: political not metaphysical", *Princeton University Press*, pp. 223-251, source Philosophy and Public Affairs, 1985.

Rawls, John, *Lecciones sobre la historia de la filosofía política*, Paidós, Barcelona, 2009.

Ribotta, Silvina, *John Rawls: sobre la (des)igualdad y justicia*, Dykinson, Madrid, 2009.

Ricoeur, Paul, *Lo justo*, Caparrós, Madrid, 2003.

Rivera, Faviola, "Rawls, filosofía y tolerancia" en *Isonomía* No. 19, UNAM, México, 2003.

Sandel, Michael, *Justice: what is the right thing to do?* Farrar Straus and Giroux, New York, 2009.

Seleme, Hugo Omar, "¿Puede el utilitarismo ser deontológico? Una respuesta a Kymlicka", *CRITICA: Revista Hispanoamericana de Filosofía*. Vol. 36, No. 107, pp. 39-66, 2004, Tomado de https://www.jstor.org/stable/40104933?seq=1#page_scan_tab_contents

Sen, Amartya, *La idea de la justicia*, Penguin Random House, México, 2009.

Schelling, Thomas, *La estrategia del conflicto*, Tecnos, Madrid, 1964.

Sidgwick, Henry, *The methods of ethics*, Macmillan and Co., New York, 1907. Tomado de <http://www.gutenberg.org/files/46743/46743-h/46743-h.htm>, el 06 de febrero de 2020.

Smart, John Jamienson y Willianms, Bernard, *Utilitarismo: pro y contra*, Tecnos, Madrid, 1981.

Sidgwick, Henry, *The methods of Ethics*, Barnes & Noble: digital library, New York, 2011.

Schultz, Bart, *The happiness philosophers: the lives and works of the great utilitarians*, Princeton University Press, New Jersey, 2017.

Vallespín Oña, Fernando, *Nuevas teorías del contrato social: John Rawls, Robert Nozick y James Buchanan*, Alianza, Madrid, 1985.

Venezia, Luciano, “Bienes primarios versus utilidad” en *Análisis filosófico*, v. 27 n.2 Ciudad Autónoma de Buenos Aires, 2007. Tomado de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-96362007000200005#notas

Waldron, Jeremy, “John Rawls and the Social Minimum”, *Journal of Applied Philosophy*, Vol.3, No. 1 (1986), pp. 21-33. Tomado de <https://www.jstor.org/stable/24353470> (16 de junio de 2020).

Walzer, Michael, *Las esferas de la justicia: una defensa del pluralismo y la igualdad*, Fondo de Cultura Económica, México, 1993.

Wolff, Robert Paul, *Para comprender a Rawls: una reconstrucción y una crítica de la teoría de la justicia*, Fondo de Cultura Económica, México, 1977.