

# POSICIÓN ORIGINAL, EQUILIBRIO REFLEXIVO Y DEONTOLOGISMO

HUGO OMAR SELEME

CONICET - Universidad Nacional de Córdoba

## RESUMEN

Un aspecto poco estudiado del argumento de Michael Sandel en contra del carácter *neutral* de la justicia como equidad, es el modo en que funda sus conclusiones en el entendimiento que tiene de otros tres aspectos de la concepción rawlsiana de justicia: su carácter *deontológico*, el *equilibrio reflexivo*, y la *posición original*. Nuestro objetivo es mostrar que Sandel no ha cometido cuatro errores independientes, sino que poseer un entendimiento equivocado del carácter *deontológico* de la teoría lo ha llevado a caracterizar mal en qué consiste el *equilibrio reflexivo* y esto –finalmente– lo ha conducido a una errónea interpretación de lo que se encuentra modelado en la *posición original* y –en última instancia– a sostener la falta de *neutralidad* de la concepción de justicia.

Concentrarnos en este asunto nos permitirá apreciar cómo se articulan y complementan las distintas réplicas que Brian Barry, Samuel Freeman y el mismo John Rawls han dirigido contra Sandel.

PALABRAS CLAVE: neutralidad, Sandel, Rawls.

## ABSTRACT

One aspect of Michael Sandel's argument against the *neutral* character of justice as fairness on which little research has been done is the way in which he founds his conclusions on the basis of the understanding he has about three other aspects of the Rawlsian conception of justice: its *deontological* character, the *reflective equilibrium* and the *original position*. It is my aim to prove that Sandel has not made four independent mistakes but, on the contrary, to have a wrong understanding of the *deontological* character of the theory has driven him to a wrong characterization of what the *reflective equilibrium* consists of and this, in turn, has taken him to a wrong interpretation of what is modelled in the *original position* and, at large, to hold the idea that justice as fairness lacks *neutrality*.

Focusing on this issue will allow us to appreciate how the different responses made by Brian Barry, Samuel Freeman and John Rawls himself against Sandel are articulated and complement one another.

KEYWORDS: neutrality, Sandel, Rawls.

## Introducción

Dado el carácter pluralista de las sociedades modernas, signadas por la existencia de individuos que profesan divergentes y hasta antagónicas concepciones del bien, una de las principales virtudes que *A Theory of Justice* (Rawls, 1971) se preció de tener desde su aparición, fue su neutralidad respecto de estas. John Rawls afirmaba que ni la elaboración de la concepción de justicia presente en esta obra, ni la aceptación de sus principios, implicaba comprometerse con alguna concepción de la vida buena.

Este objetivo era logrado gracias al carácter *deontológico* de la teoría, esto es, gracias a su consideración de que lo correcto era prioritario en relación con lo bueno. Sin embargo, no fue este el único aspecto de la concepción de justicia que atrajo a los teóricos. En particular otros dos concitaron su atención: la utilización del recurso de la *posición original* como herramienta de construcción de los principios de justicia, y la utilización del *equilibrio reflexivo* como novedoso<sup>1</sup> método de justificación y evaluación.

Uno de los asuntos en que se ha centrado el interés académico ha sido el referido a si la utilización conjunta de estos dos recursos, *la posición original* y *el equilibrio reflexivo*, herederos de tradiciones tan disímiles como el contractualismo y el coherentismo, es posible<sup>2</sup>. Un tópico menos analizado ha sido el referido a la vinculación que existe entre la forma de entender *la prioridad de lo correcto*, la manera de interpretar el *equilibrio reflexivo* y el diseño de *la posición original*. Es decir, ha existido un extenso debate sobre cada uno de estos temas en particular, pero no sobre las implicancias que tienen en los otros asuntos las posiciones que se toman en relación con cada uno de ellos.

Uno de los pocos que acometió la tarea de vincularlos ha sido Michael Sandel en *Liberalism and the Limits of Justice* (Sandel,

<sup>1</sup> El método era novedoso, no porque Rawls fuese el primero en utilizarlo para justificar principios, sino porque era el primero en aplicarlo a la justificación de principios de moralidad política. Rawls mismo señala en *A Theory of Justice* que Nelson Goodman ya lo había utilizado en *Fact, Fiction, and Forecast* (Cambridge, Mass., Harvard University Press, 1955) para justificar los principios de inferencia deductiva e inductiva (Rawls, 1971 p. 18, nota 7).

<sup>2</sup> Cf. Barry (1989), Dworkin (1975), Lyons (1975) y Daniels (2000). Para una defensa de la utilización conjunta de ambos recursos –el contractualista, corporizado en *la posición original*, y el coherentista, corporizado en el *equilibrio reflexivo*– véase Seleme (2004).

1982). Su forma de interrelacionarlos ha tenido gran relevancia, además, porque Sandel ha construido, basado en ella, un poderoso argumento con objeto de mostrar que la justicia como equidad carece de neutralidad metafísica, es decir que se encuentra comprometida con una particular concepción metafísica de persona, tan controvertida como cualquier concepción de la vida buena.

Si el argumento de Sandel es correcto, de nada le valdría a la concepción de justicia preservar su neutralidad en relación con las concepciones del bien, ya que no sería neutral en un ámbito —el metafísico— en el cual los individuos de una sociedad democrática y pluralista mantienen tantas discrepancias como las que sostienen respecto a sus concepciones de la vida buena. Así, una de las principales virtudes de la justicia como equidad, esto es, que podría ser aceptada por los individuos de una sociedad pluralista, se habría perdido. Los ciudadanos que la aceptaran deberían adherir todos a una misma concepción metafísica de la persona humana.

El argumento de Sandel para mostrar este compromiso metafísico de la concepción rawlsiana de justicia hace hincapié en una particular interpretación del diseño de la *posición original*. El entramado argumentativo en el que funda su interpretación es de gran complejidad e involucra no solo una manera de entender el recurso del *equilibrio reflexivo*, sino también una forma particular de comprender la exigencia deontológica de dar *prioridad a lo correcto en relación con lo bueno*. Cuatro tópicos fundamentales de la justicia como equidad han sido vinculados por Sandel: su *neutralidad*, su carácter *deontológico*, su utilización del *equilibrio reflexivo* como método de justificación, y su diseño de *la posición original*.

Las posiciones que Sandel ha tomado en relación con cada uno de estos asuntos han sido cuestionadas tanto por el propio Rawls —quien desde la publicación de “Justice as Fairness: Political not Metaphysical” (Rawls, 1985) ha venido llamando la atención sobre el hecho de que la concepción de persona modelada en la *posición original* es de tipo normativo y no metafísico— como por otros autores tales como Samuel Freeman (Freeman, 1994) y Brian Barry (Barry, 1995). Sin embargo se ha prestado menos atención a la manera en que las consideraciones de Sandel se encuentran vinculadas entre sí, es decir, al modo en que lo sostenido respecto a cada una de ellas influye en las otras. El presente trabajo está dirigido a remediar esta falencia.

Nuestro objetivo es mostrar la manera incorrecta en que Sandel correlaciona estos cuatro asuntos (falta de *neutralidad*, carácter

*deontológico* de la teoría, *equilibrio reflexivo* y *posición original*) y no solo el modo en que sus posiciones en relación con cada uno de ellos son erróneas. Sandel no ha cometido cuatro errores independientes, sino que poseer un entendimiento equivocado del carácter *deontológico* de la teoría lo ha llevado a caracterizar mal en qué consiste el *equilibrio reflexivo* y esto, finalmente, lo ha conducido a una errónea interpretación de lo que se encuentra modelado en la *posición original* y, en última instancia, a sostener la falta de *neutralidad* de la concepción de justicia.

Concentrarnos en la manera en que Sandel conecta la interpretación de los distintos elementos tiene dos ventajas: en primer lugar, al mostrar una forma de vinculación errónea, nos permite avanzar en la tarea de dilucidar cuál es la verdadera interrelación que existe entre estos temas; en segundo lugar, nos posibilita articular y ver cómo se complementan las distintas réplicas que autores como Barry, Freeman y el mismo Rawls han dirigido contra el argumento de Sandel.

El presente trabajo está dividido en cuatro secciones. La primera está dedicada a esclarecer la particular manera en que Sandel interpreta la *prioridad* que la justicia como equidad otorga a *lo correcto en relación con lo bueno*. En la segunda el tema es su entendimiento del *equilibrio reflexivo*. En la tercera sección se intenta mostrar de qué manera las posturas que Sandel asume respecto a los dos tópicos anteriores le permiten ver modelada en la *posición original* una particular concepción metafísica de persona. En la última sección, se presenta la particular concepción metafísica que Sandel encuentra modelada en los elementos de la *posición original*, que compromete la *neutralidad* de la concepción de justicia. Finalmente, una vez que se han mostrado las vinculaciones existentes entre las distintas posiciones adoptadas por Sandel, se articulan en una misma línea argumental las réplicas formuladas por Barry, Freeman y el propio Rawls.

## 1. Prioridad de lo correcto en relación con lo bueno

Como hemos señalado parte del objetivo del presente trabajo es mostrar de qué forma la manera en que Sandel interpreta la *posición original* está determinada por su particular entendimiento del carácter *deontológico* de la concepción de justicia rawlsiana. Para lograrlo es necesario, en primer lugar, analizar el modo en que

Sandel interpreta el hecho de que una concepción de justicia posea este carácter.

Según él, una de las características de las concepciones de justicia liberales es la *primacía o prioridad que otorgan a la justicia*. Esta primacía puede ser entendida de dos maneras diferentes: como una forma de prioridad moral o como una forma privilegiada de justificación. Entendida en su sentido moral, la primacía se refiere a que las exigencias de justicia, y los derechos individuales que de ellas se derivan, tienen más peso que otros intereses morales o políticos<sup>3</sup>. En su segundo sentido, la primacía se refiere a que los principios de justicia tienen una forma privilegiada de justificación. La primacía de lo justo sobre otros valores morales o políticos, o la primacía de lo correcto sobre lo bueno, no se refiere al mayor peso o importancia moral o motivacional, sino a que los principios de justicia no están justificados basándose en ninguna concepción particular del bien<sup>4</sup>.

Para Sandel, el liberalismo *deontológico* se diferencia del *teleológico* en que este último acepta la primacía moral de la justicia, pero no su modo privilegiado de justificación, mientras que el primero sostiene que la primacía moral no puede ser garantizada a menos que se justifique a los principios de justicia de una manera privilegiada. Así, el liberalismo *deontológico* –a diferencia del *teleológico*– busca obtener principios de justicia de un modo que no presuponga ningún propósito o fin humano o, lo que es lo mismo, ninguna concepción de la vida buena.

Ambos tipos de liberalismo, argumenta Sandel, afirman que las exigencias fundadas en la justicia poseen una precedencia categórica respecto a otras exigencias morales, aunque discrepan en

<sup>3</sup> Señala Sandel: “On this view, justice is not merely one value among others, to be weighed and considered as the occasion arises, but the highest of all social virtues, the one that must be met before others can make their claims. If the happiness of the world could be advanced by unjust means alone, not happiness but justice would properly prevail. And when justice issues in certain individual rights, even the general welfare cannot override them” (Sandel, 1982, p. 2). Sandel habla aquí a la vez de “peso” como sinónimo de importancia moral y de fuerza motivacional. En consecuencia en lo sucesivo cuando hablemos de “primacía moral” deberá entenderse como Sandel lo hace, es decir, como primacía en importancia moral y en fuerza motivacional.

<sup>4</sup> Sostiene al respecto: “Unlike other practical injunctions, principles of justice are justified in a way that does not depend on any particular vision of the good. To the contrary given its independent status, the right constrains the good and sets its bounds” (Sandel, 1982, p. 2).

cuál es el modo en que los principios de justicia deberían justificarse<sup>5</sup>. El liberalismo *deontológico* reclama, a diferencia del liberalismo *teleológico*, que ambos sentidos de la primacía de la justicia no pueden entenderse de manera separada. La primacía en sentido moral no puede justificarse ni explicarse si no se acepta la primacía en sentido justificatorio. Esto debido a que si reconocemos el fundamento último de la primacía moral de la justicia en su contribución a un objetivo que se considera valioso –tal como sostiene el liberalismo *teleológico*–, tendríamos que admitir que aquella primacía es contingente. Puestos en circunstancias en las que actuar injustamente maximizara aquel valor, no tendríamos razón alguna para darle prioridad a las exigencias derivadas de la justicia (Sandel, 1982, p. 4). Así, si se pretende garantizar la primacía en sentido moral de los principios de justicia, sostiene el liberalismo *deontológico*, en la derivación de estos no debe apelarse a ninguna concepción particular del valor o de la vida buena.

Con estas precisiones a mano, concluye Sandel, es posible advertir que el liberalismo corporizado en la justicia como equidad es un tipo de *deontologismo*. Le otorga a la justicia una prioridad absoluta en sentido moral<sup>6</sup> y le reconoce un modo privilegiado de justificación que no toma como base ninguna concepción del bien<sup>7</sup>.

<sup>5</sup> En relación con esto afirma: “The contrast might also be drawn in terms of two different senses of deontology. In its moral sense, deontology opposes *consequentialism*; it describes a first-order ethic containing certain categorical duties and prohibitions which take unqualified precedence over other moral and practical concerns. In its foundational sense, deontology opposes *teleology*; it describes a form of justification in which first principles are derived in a way that does not presuppose any final human purposes or ends, nor any determinate conception of the human good” (Sandel, 1982, p. 3).

<sup>6</sup> Sandel, para justificar su posición, hace referencia a la siguiente afirmación de Rawls: “Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust [...]. Being first virtues of human activities, truth and justice are uncompromising” (Rawls, 1971, pp. 3-4). Véase Sandel, 1982, p. 15.

<sup>7</sup> Para dar sustento a su afirmación Sandel cita a Rawls señalando: “the structure of teleological doctrines is radically misconceived: from the start they relate the right and the good in the wrong way. We should not attempt to give form to our life by first looking to the good independently defined. It is not our aims that primarily reveal our nature but rather the principles that we would acknowledge to govern the background conditions under which these aims are to be formed and the manner in which they are to be pursued. *For the self is prior to the ends which are affirmed by it*; even a dominant end must be chosen from among numerous possibili-

## 2. El recurso del equilibrio reflexivo

Habiendo aclarado cómo Sandel interpreta el carácter deontológico de la justicia como equidad y su exigencia de otorgar *prioridad a lo correcto en relación con lo bueno*, es necesario exponer el segundo conjunto de consideraciones que dan sustento a su interpretación de la *posición original*: aquellas referidas a su manera de entender el *equilibrio reflexivo*.

Para Sandel, el método del *equilibrio reflexivo* busca, por un lado, articular en una concepción de justicia nuestras convicciones sobre lo que consideramos correcto y, por otro, articular una concepción metafísica de persona a la cual pueda aplicarse aquella concepción de justicia. El procedimiento para alcanzar el equilibrio nos exige movernos entre los dos extremos, por una parte, configurando los principios de justicia de manera que se ajusten con nuestros juicios meditados de moralidad política, y por la otra, configurando una concepción de persona que sea plausible, que tenga características semejantes a las de un individuo humano, para el cual aquellos principios y la concepción de la que forman parte puedan ser los adecuados (Sandel, 1982, p. 48).

Ahora bien, como la exigencia de prioridad o primacía moral es parte de la concepción de justicia, únicamente se ha alcanzado el *equilibrio reflexivo* cuando han sido obtenidos los principios más convenientes desde el punto de vista moral, que pueden ser la *primera virtud* para una persona concebida de una forma determinada, concepción que a su vez nos parece plausible por estar configurada con rasgos que razonablemente atribuiríamos a un individuo humano.

De este modo, concluye Sandel, dado que Rawls utiliza como método de justificación de su teoría el *equilibrio reflexivo*, el producto final al que arriba tiene que ser no solo una teoría de la justicia sino también una concepción de persona para la cual las exigencias de justicia deban tener primacía moral.

---

lities [...]. We should therefore reverse the relation between the right and the good proposed by teleological doctrines and view the right as prior. The moral Theory is then developed by working in the opposite direction" (Rawls, 1971, pp. 490-491). Véase Sandel, 1982, p. 19. (Las bastardillas pertenecen a la cita de Sandel.)

### 3. La posición original como modeladora de una concepción metafísica de persona

Teniendo a mano la interpretación que Sandel brinda del *equilibrio reflexivo* y la *prioridad de lo correcto*, es posible ya abordar la particular manera en que entiende la *posición original*. Su interpretación sostiene, en primer lugar, que la *posición original* es una herramienta para alcanzar el *equilibrio reflexivo*, y que, como tal, modela los distintos elementos que buscan *equilibrarse*. En segundo lugar, afirma que estos elementos son de dos clases: a) nuestras intuiciones con relación a lo justo, y b) nuestra manera de concebir a un sujeto que se asemeje a un individuo humano y para el cual las exigencias de justicia tengan primacía moral.

El resultado del procedimiento de equilibrio alcanzado mediante la *posición original* tiene, entonces, que ser doble: a) una concepción de justicia, que se encuentre en equilibrio con b) una concepción de persona.

Hasta aquí nada se ha dicho en relación con el carácter metafísico de esta concepción de persona. ¿Por qué Sandel piensa que la concepción de persona necesariamente debe ser de índole metafísica? Su argumento para justificar esta aseveración puede ser reconstruido como un silogismo disyuntivo cuya primera premisa sostiene que aquellos que modelan los elementos de la *posición original* en relación con la persona son o bien generalizaciones empíricas, o bien consideraciones metafísicas. El paso siguiente es mostrar que, si se acepta uno de los extremos de la disyunción —el que se refiere a la generalización empírica—, la concepción de persona no puede hallarse en *equilibrio* con la concepción de justicia. Se obtiene así la conclusión de que aquello que subyace detrás de la *posición original* —como herramienta para alcanzar el *equilibrio reflexivo*— debe ser una concepción metafísica de persona.

El punto sobre el cual el argumento de Sandel hace hincapié, y que da sustento a su primera premisa, está vinculado con su manera de entender el *equilibrio reflexivo*. Las condiciones corporizadas en la *posición original* no solo deben ser vistas como adecuadas de acuerdo a nuestros juicios meditados respecto de la justicia, sino que tienen que guardar alguna relación con la situación en que se encuentran los seres humanos, de manera que los principios obtenidos por medio de ella puedan ser principios que se les apliquen. Es justamente esto último —el que los principios de justicia no parezcan tener ninguna relación con seres que puedan considerarse semejan-



tes a hombres— lo que Rawls, en opinión de Sandel, ve como inadecuado en la doctrina kantiana (Sandel, 1982, p. 43).

La cuestión entonces es, según Sandel, de qué manera Rawls pretende que las condiciones corporizadas en la *posición original* tengan algún parecido con la situación en que se encuentran criaturas discerniblemente humanas, para quienes los principios de justicia puedan tener la prioridad moral que su teoría les otorga. ¿Cómo lograr que la *posición original* modele una concepción de persona plausible, para la cual las exigencias de justicia tengan prioridad moral categórica<sup>8</sup>?

Las alternativas, señala, son dos: o bien las condiciones descritas en la *posición original* deben tomarse como asentadas en generalizaciones empíricas —de tipo sociológico— con relación a cuáles son en general los motivos de los seres humanos y cuáles son las condiciones que prevalecen en sus organizaciones sociales; o bien deben tomarse como elementos que no intentan describir cómo *de hecho* se comportan los seres humanos, sino que son parte de una situación hipotética que intenta modelar una concepción metafísica de persona.

Establecida la primera premisa del silogismo disyuntivo, el próximo paso es analizar sus alternativas. Dentro de los elementos de la *posición original* Sandel presta especial atención a las *circunstancias de la justicia* (Sandel, 1982, p. 29). Refiriéndose a ellas, indica Rawls:

For simplicity I often stress the condition of moderate scarcity (among the objective circumstances), and that of conflict of interests (among the subjective circumstances). Thus, one can say, in brief, that the circumstances of justice obtain whenever mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity. Unless these circumstances existed there would be no occasion for the virtue of justice, just as in the absence of threats of injury to life and limb there would be no occasion for physical courage (Rawls, 1971, p. 128)<sup>9</sup>.

<sup>8</sup> Señala Sandel: “assuming we are beings capable of justice, and more precisely, beings for whom justice is primary, we must be creatures of a certain kind, related to human circumstances in a certain way. What then must be true of a subject for whom justice is the first virtue? And how is the conception of such a subject embodied in the original position?” (Sandel, 1982, p. 49).

<sup>9</sup> A diferencia del resto de las citas de *A Theory of Justice*, esta no ha sido tomada de la edición revisada, sino de la original. La razón es que, sugestivamente, Rawls ha modificado este pasaje en la última edición de *A Theory of Justice*. En esta

La primera manera de entender las *circunstancias de la justicia*, señala Sandel, es verlas como generalizaciones empíricas respecto de los individuos y las sociedades humanas. Entendidas de este modo, Rawls estaría afirmando que –puesto que las sociedades humanas están caracterizadas por estas circunstancias empíricas– la virtud que se considera más valiosa respecto de sus instituciones sociales es aquella que intenta enfrentar los problemas a los cuales estas circunstancias dan lugar, en especial, los referidos al reparto de los beneficios y cargas producto de la cooperación social entre seres que poseen distintos planes de vida y, por lo tanto, sostienen posiciones divergentes respecto a cómo debería ser la distribución.

Sin embargo, señala Sandel, si se da esta interpretación empírica a las *circunstancias de la justicia*, la concepción de persona, aunque plausible, no logra garantizar la primacía moral de las exigencias fundadas en la justicia. La concepción de persona no se encuentra en equilibrio con la concepción de justicia que sostiene aquel carácter prioritario de sus exigencias. El carácter prioritario de la justicia se encuentra amenazado de dos modos.

a) Como el valor de la justicia como virtud está condicionado a que se den ciertas circunstancias empíricas, su primacía –en sentido moral– no podría afirmarse incondicionalmente como pretende, dado el carácter deontológico de su doctrina, Rawls.

Una manera posible en que Rawls podría sostener la *primacía de la justicia* –en sentido moral– y a la vez utilizar en la *posición original* las *circunstancias de la justicia* –en su interpretación empírica– sería mostrar, mediante argumentos de tipo sociológico, que

---

se ha suprimido la referencia al desinterés mutuo de las personas: “For simplicity I often stress the condition of moderate scarcity (among the objective circumstances), and that of conflict of interests (among the subjective circumstances). Thus, one can say, in brief, that the circumstances of justice obtain whenever persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity. Unless these circumstances existed there would be no occasion for the virtue of justice, just as in the absence of threats of injury to life and limb there would be no occasion for physical courage” (edición revisada, p. 110). Pensamos que la razón de la modificación debe buscarse en el hecho de que si la teoría pretende permanecer neutral en relación con las concepciones del bien profesadas por las personas, entonces no puede sostener que estas son mutuamente desinteresadas. Las circunstancias de la justicia, como quedará claro más adelante, hacen referencia a la existencia de reclamos conflictivos basados en intereses dignos de consideración, no en las motivaciones de estos reclamos. El mutuo desinterés es una característica motivacional de las partes en la *posición original*, que tiene por objeto modelar el hecho de que las personas realizan reclamos conflictivos por la división de recursos sociales escasos basados en intereses que consideran dignos de ser tomados en cuenta.

en las sociedades industriales modernas esas condiciones prevalecen de tal modo que hacen que la justicia tenga la primacía moral que le asigna (Sandel, 1982, p. 30).

No obstante, aun en este supuesto, si se adopta la interpretación empirista de las *circunstancias de la justicia* no podría hacerse la afirmación categórica que Rawls efectúa en el sentido de que la justicia es la virtud más importante de las instituciones sociales; sino solo la afirmación hipotética de que si en una sociedad prevalecen las circunstancias de la justicia y los individuos se comportan por lo general como estas requieren, entonces allí se considerará a la justicia como la principal virtud de sus instituciones sociales<sup>10</sup>.

Por lo tanto, en opinión de Sandel, adoptar esta interpretación de la *posición original* impide adjudicarle a los principios de justicia la primacía en sentido moral que las teorías deontológicas –incluida la de Rawls– buscan asegurar concediendo *prioridad* –en sentido justificativo– *a lo correcto en relación con lo bueno*.

b) El segundo modo en que la primacía se vería amenazada sería el siguiente. Mientras que en el caso anterior se señala que sostener una interpretación empírica de las circunstancias de la justicia socava la *primacía de la justicia*, por volver a sus exigencias condicionales antes que categóricas, esta segunda objeción sostiene que aquella interpretación pone en riesgo esta primacía de la justicia, por equipararla en importancia con virtudes que se dan en otras situaciones a las que debería reconocérseles un valor idéntico.

Aquí no se pone el acento en el carácter hipotético de las exigencias de justicia, sino en su carácter *reparador*. Dado que la justicia es considerada valiosa únicamente en aquellas situaciones donde prevalecen las *circunstancias de la justicia* –puesto que es vista como un remedio a las dificultades que estas circunstancias generan–, el que aquellas circunstancias no prevalezcan en una determinada sociedad debe considerarse algo tan valioso como el hecho de que sus instituciones sean justas. En consecuencia, no sería correcto afirmar que la justicia es el principal valor de las instituciones sociales. Una sociedad en la que no se dieran las *circunstan-*

<sup>10</sup> Refiriéndose a esta primera objeción fundada en la interpretación empirista de las “circunstancias de la justicia”, afirma Sandel: “On the empirist interpretation of the original position, justice can be primary only for those societies beset by sufficient discord to make the accommodation of conflicting interests and aims the overriding moral and political consideration; justice is the first virtue of social institutions not absolutely, as truth is to theories, but only conditionally, as physical courage is to a war zone” (Sandel, 1982, p. 31).

*cias de la justicia* sería tan valiosa como una sociedad en la cual sus instituciones fueran justas<sup>11</sup>.

De lo dicho se sigue que si la *posición original* es una herramienta para alcanzar el *equilibrio reflexivo* entre la concepción de justicia que nos parece correcta –en relación con nuestros juicios morales meditados– y la concepción de persona para la cual aquellos criterios de justicia son adecuados, y si parte de la concepción de justicia es el considerar que las exigencias de justicia son *prioritarias* –tienen *primacía moral y motivacional* sobre otras consideraciones–, entonces las premisas descriptivas de la *posición original* no pueden estar fundadas en contingentes generalizaciones empíricas. La concepción de persona modelada en la *posición original* no puede tener carácter empírico.

Así, habiendo descartado la primera cláusula de la premisa disyuntiva, Sandel se queda con la segunda: la *posición original* es una situación hipotética y las condiciones impuestas en ella no deben ser tomadas como las condiciones en las que en realidad viven los seres humanos, ni como las condiciones en las que se encuentran las sociedades en las que ellos habitan<sup>12</sup>. El criterio para evaluar la plausibilidad de las premisas descriptivas de la *posición original* no son generalizaciones empíricas sino, afirma Sandel, los juicios que tenemos en relación con la manera de concebir al sujeto moral, es decir, el entendimiento que tenemos de nosotros mismos.

Esta manera de entender el criterio de plausibilidad de las premisas descriptivas de la *posición original* no apela a generalizaciones empíricas, ni a nuestras “intuiciones” sobre lo que consideramos empíricamente verdadero acerca del ser humano, sino a la

<sup>11</sup> Afirma Sandel en relación con lo señalado: “But if the virtue of justice is measured by the morally diminished conditions that are its prerequisite, then the *absence* of these conditions [...] must embody a rival virtue of at least commensurate priority, the one that is engaged in so far as justice is not engaged” (Sandel, 1982, p. 32). Posteriormente, Sandel extiende su argumento para sostener que no solo la justicia debe considerarse como un valor equivalente al corporizado en aquellas situaciones donde las circunstancias de la justicia no prevalecen, sino como un valor de menor jerarquía. Al respecto señala: “An increase in justice can fail to be associated with an overall moral improvement in at least two different ways: either by failing fully to meet an increase in the circumstances of justice, or by an inability, however full, to compensate the loss of certain ‘nobler virtues, and more favourable blessings’” (Sandel, 1982, p. 34).

<sup>12</sup> Afirma al respecto: “The conditions described there [en la *posición original*] are meant to be the conditions in which the parties to the original position carry out their deliberations, not the actual conditions in which ordinary human beings live their lives” (lo agregado entre corchetes nos pertenece) (Sandel, 1982, p. 41).

particular manera que tenemos de vernos a nosotros mismos como sujetos morales para quienes es adecuada la concepción de justicia, es decir, la comprensión que tenemos de nuestra naturaleza profunda, no de nuestros accidentes empíricos por más regulares que estos puedan ser.

El afirmar que las premisas descriptivas de la *posición original* modelan una concepción antropológica de tipo metafísica, y no una mera generalización empírica, permite que la concepción de persona corporizada en ella pueda encontrarse en *equilibrio* con la concepción de justicia que adjudica a sus exigencias primacía moral y motivacional<sup>13</sup>. En efecto, la primacía de los principios de justicia depende de una particular manera de entender al sujeto moral, pero este entendimiento no se obtiene mediante generalizaciones empíricas, sino mediante autorreflexión. Esto le da a la concepción un carácter menos contingente, en el sentido de que el sujeto moral es lo que actualmente es, en parte debido a su manera de concebirse a sí mismo. Es decir, si no fuera capaz de percibirse a sí mismo como tal sujeto moral, entonces no lo sería. De esta forma, una concepción antropológica –de índole metafísica– no es algo que simplemente describe atributos contingentes del sujeto, sino que en cierto modo lo constituye como tal<sup>14</sup>.

Con todas las precisiones formuladas a mano, pensamos que es posible ahora advertir con toda claridad la manera en que las distintas posiciones de Sandel se encuentran interrelacionadas.

<sup>13</sup> Exponiendo su manera particular de entender el *equilibrio reflexivo* y su utilización para modelar la *posición original*, afirma Sandel: “To summarize: the validity of a premise of the original position is not given empirically, but by a method of justification known as reflective equilibrium. This method involves two different kinds of justification coming together to provide mutual correction and support. One aspect of the justification appeals to our considered convictions about justice; the other appeals to a standard of descriptive but not empirical plausibility” (Sandel, 1982, p. 44).

<sup>14</sup> Dice Sandel al respecto: “Now the description of this subject will have a distinctive logical status. It will in some sense be necessary, non-contingent, and prior to any particular experience [...] but it will not of course be an analytic claim. It will in some sense be empirical, but not ‘merely’ empirical. Given the reflexive character of such descriptions, they are not merely descriptive but also partly constitutive of the kind of beings we are. Our knowing them is part of what makes them true, and makes us the reflexive, self-interpreting creatures that we are [...]. To avoid confusion on this count, I should say in advance that the account I have in mind is a *philosophical anthropology* in the best sense; *philosophical* in that it is arrived at reflectively rather than by empirical generalization, *anthropology* in that it concerns the nature of the human subject in its various possible forms of identity” (las cursivas son nuestras) (Sandel, 1982, p. 49-50).

En primer lugar, Sandel interpreta que la concepción de justicia rawlsiana afirma la primacía –moral y motivacional– de las exigencias de justicia. Este es uno de los sentidos en que lo justo es *prioritario en relación con lo bueno*. Las exigencias fundadas en los principios de justicia nunca pueden ser superadas por las exigencias fundadas en una concepción de la vida buena. A su vez, para garantizar esta primacía es necesario que los principios de justicia no se encuentren fundados en una particular concepción del bien.

En segundo lugar, Sandel entiende que el método del *equilibrio reflexivo* busca articular la concepción de justicia que nos parece más deseable con una concepción de persona que nos parece plausible. Dado que uno de los elementos de la concepción de justicia es la primacía moral de sus principios, debe buscarse una concepción plausible de persona para la cual las exigencias de justicia tengan de modo necesario esta primacía.

Finalmente, la *posición original* es una herramienta para alcanzar el *equilibrio reflexivo* y, como tal, debe modelar una concepción de persona para quien las exigencias de la concepción de justicia sean adecuadas. Ahora bien, dado que una de las características de la concepción de justicia rawlsiana es el otorgar *primacía categórica o incondicional a las exigencias de justicia*, las premisas descriptivas de la *posición original* deben modelar una concepción de persona para la cual las exigencias de la justicia tengan esta prioridad de manera categórica y no meramente contingente. Una de las exigencias para satisfacer este requisito es que la concepción de persona misma sea necesaria y no contingente. Si la concepción de persona modelada en la *posición original* tiene carácter metafísico, y no empírico, esta exigencia ha sido satisfecha y la primacía de la justicia puede ser garantizada. Es decir, mientras los seres humanos sigan siendo los sujetos autorreflexivos que son, los principios de justicia –que se derivan de esta concepción de persona– seguirán teniendo primacía.

Habiendo vinculado la interpretación que Sandel brinda del carácter *deontológico* de la concepción rawlsiana con su entendimiento del *equilibrio reflexivo* y de la *posición original*, resta todavía por mostrar de qué modo esto se vincula con su posición en relación con la falta de *neutralidad* de la justicia como equidad. Para ello presentaremos brevemente cuál es la concepción metafísica que Sandel considera subyacente al diseño de la *posición original* y la razón por la que piensa que compromete el carácter neutral de la justicia como equidad.

## 4. La concepción metafísica de persona y la falta de neutralidad de la concepción de justicia

### 4.1. Las circunstancias de la justicia y la prioridad de la pluralidad

Según Sandel, lo que se encuentra detrás de la caracterización de las *circunstancias de la justicia* –y los problemas a los que estas dan lugar– es una manera particular de concebir la sociedad como una empresa cooperativa para beneficio mutuo, caracterizada a la vez por el conflicto y la identidad de intereses. Esta manera de entender la sociedad y el problema al que intenta dar solución la justicia implican una manera de concebir al ser humano. Así, si la justicia va a ser la primera virtud –y a tener primacía moral–, dado que esta únicamente tiene lugar cuando se presentan intereses en conflicto, los seres humanos ante todo deben ser diferentes unos de otros, esto es, deben ser plurales<sup>15</sup>.

Rawls señala que la sociedad como empresa cooperativa no solo está caracterizada por el conflicto, sino también por la identidad de intereses. Es decir, para que la cooperación sea posible tiene que afirmarse tanto la pluralidad como la unidad de sus miembros. Ahora bien, según Sandel, a estos dos hechos –que cada individuo sea alguien distinto entre una multiplicidad de seres y que sus intereses y propósitos puedan coincidir con los de otro– Rawls les otorga distinta importancia. Esto porque para que un sujeto sea capaz de practicar la justicia, lo que se necesita es que tenga intereses divergentes con los demás; el hecho de que sus intereses converjan es solo una feliz coincidencia que hace a la cooperación ventajosa<sup>16</sup>.

En consecuencia, lo que se encuentra detrás de la utilización de las *circunstancias de la justicia* es –por un lado– una concepción

<sup>15</sup> Sostiene Sandel: “For there to be justice, there must be the possibility of conflicting claims, and for there to be conflicting claims, there must be more than a single claimant. In this way, the plurality of persons can be seen as a necessary presupposition of the possibility of justice” (Sandel, 1982, p. 50-51).

<sup>16</sup> Señala al respecto Sandel: “the conflict of interests arises, as we have seen, from the fact that the subjects of co-operation have different interests and ends, and *this* fact follows from the nature of a being capable of justice at all. The identity of interests, however, expresses the fact that the parties happen to have suitably similar needs and interests such that co-operation among them is mutually advantageous. And *this* fact, that their needs and interests happen to coincide in such a way, does not follow from the nature of their subjectivity, but merely from the happy accident of their circumstances” (Sandel, 1982, p. 53).

de la sociedad entendida como una empresa cooperativa para la ventaja mutua y –por el otro– una concepción del sujeto moral como plural<sup>17</sup>, esto es, caracterizado por un sistema separado de fines. Basta saber que los individuos participan en la sociedad entendida de esta manera y que son seres capaces de practicar la justicia, para poder afirmar que poseen un sistema separado de fines que los distingue.

Para que la justicia tenga la *primacía* que la teoría de Rawls le adjudica, concluye Sandel, la pluralidad de las personas tiene que ser un aspecto fundamental de la naturaleza humana. Los sujetos primero tienen que estar individualizados, ser unos distintos de otros por poseer diferentes sistemas de fines, y recién luego pueden establecer relaciones con los demás. La identidad del sujeto debe estar constituida con anterioridad a los vínculos que lo ligan con otros. La pluralidad es prioritaria en relación con la unidad, y los vínculos que nos unen a nuestros pares nunca pueden constituir nuestra individualidad.

De esta manera, para Sandel, el concebir a la sociedad como un sistema de cooperación implica concebir a los sujetos que participan de ella como constituidos con anterioridad a su ingreso en la empresa cooperativa, en la que toman parte únicamente con la finalidad de buscar su ventaja mutua. Su pluralidad, el ser distintos unos de otros, es anterior a su unidad, es decir, a su participación en una empresa compartida<sup>18</sup>.

## 4.2. El supuesto del mutuo desinterés y el sujeto de posesión

Un punto que a Sandel le parece enigmático en la presentación que Rawls hace de la *posición original* es que este, por un lado,

<sup>17</sup> La expresión está utilizada aquí como sinónimo de “pluralidad de sujetos morales”.

<sup>18</sup> Sandel formula dos argumentos más para abonar su particular interpretación de la *posición original* como fundada en una concepción de la sociedad, y esta en una concepción antropológica del sujeto como plural, o sea, como anteriormente individualizado. El primero hace hincapié en el contractualismo rawlsiano. Sandel sostiene que detrás de la estrategia contractual se encuentra la idea de concebir a los individuos como distintos y como individualizados con anterioridad a sus relaciones sociales. Para sustentar su argumento cita la afirmación de Rawls en el sentido que “principles of justice deal with conflicting claims upon the advantages won by



considera que el principal supuesto motivacional de las partes es el *mutuo desinterés* y, por el otro, afirma que el diseño de la *posición original* no implica comprometerse con ninguna teoría acerca de cómo son las motivaciones humanas básicas (Sandel, 1982, p. 54)<sup>19</sup>.

Lo que Sandel propone, para resolver esta aparente paradoja, es interpretar que el supuesto del *mutuo desinterés* de las partes no representa una determinada teoría en relación con los motivos humanos sino una teoría en relación con el sujeto de esos motivos, esto es, una teoría antropológica de índole metafísica que sostiene que, sean cuales sean los intereses de los individuos –egoístas, altruistas o de otro tipo–, siempre son intereses *pertenecientes* a alguien<sup>20</sup>.

Sandel sostiene, entonces, que la utilización en la *posición original* del supuesto del *mutuo desinterés* de las partes lo compromete a Rawls con una determinada manera de entender la relación que existe entre el sujeto y sus fines. Esta relación es un vínculo de posesión. Esta característica del sujeto, unida a la analizada con anterioridad, a saber, aquella que afirma la prioridad de su pluralidad –su individualización con anterioridad a cualquier experiencia– configura la concepción antropológica que modela la *posición original*. La concepción metafísica de persona sostiene, por lo tanto, que el sujeto está individualizado –o constituido– con anterioridad e independencia a cualquier relación social o fin que elija perseguir, y

---

social cooperation; they apply to the relations among several persons or groups. The word ‘contract’ suggest this plurality” (Rawls, 1971, p. 14). Véase Sandel, 1982, p. 50. El segundo se refiere a la crítica que Rawls formula al utilitarismo por no tomar seriamente la distinción entre personas. Señala: “There is no reason to suppose that the principles which should regulate an association of men is simply an extension of the principle of choice for one man. On the contrary: if we assume that the correct regulative principle for anything depends on the *nature of that thing, and that plurality of distinct persons with separate systems of ends is an essential feature of human societies*, we should not expect the principles of social choice to be utilitarians” (Rawls, 1971, p. 25). Véase Sandel, 1982, p. 51. (Las bastardillas pertenecen a la cita de Sandel.)

<sup>19</sup> Rawls sostiene en un mismo párrafo: “I shall assume that the parties in the original position are mutually disinterested: they are not willing to have their interests sacrificed to the others. The intention is to model men’s conduct and motives in cases where questions of Justice arise [...]. The account of these conditions involves no particular Theory of human motivation” (Rawls, 1971, p. 112).

<sup>20</sup> Para fundar esta interpretación Sandel utiliza el pasaje de Rawls en el que este señala: “although the interests advanced by these plans are not assumed to be interests in the self, they are the interests of a self that regards its conception of the good as worthy of recognition” (Rawls, 1971, p. 110).

que se relaciona con estos como lo haría con cualquier objeto de su elección<sup>21</sup>.

Ahora es posible apreciar con claridad por qué esto compromete el carácter neutral de la concepción de justicia: esta concepción de persona no respeta las convicciones de aquellas personas que se ven a sí mismas como constituidas por sus lazos comunales o sociales. No respeta a aquellos individuos que, por ejemplo, no consideran a sus fines religiosos como producto de su libre elección, sino como constitutivos de los individuos que son. Dicho de modo general, esta no es más que una de las muchas posibles concepciones metafísicas que los individuos pueden profesar y la teoría de justicia fundada en ella, en consecuencia, no sería neutral con relación al resto de las concepciones metafísicas a las que podrían adherir los individuos.

## 5. Articulación de las réplicas a Sandel

El argumento ofrecido por Sandel se mueve, en primer lugar, desde una interpretación de la *prioridad de lo correcto* a un particular entendimiento del *equilibrio reflexivo*. Dado que parte de la concepción rawlsiana de justicia es su afirmación de la *prioridad de lo correcto* –interpretada por Sandel como *primacía de la justicia*–, y dado que el *equilibrio reflexivo* es una herramienta de elaboración y justificación de la concepción de justicia, lo que el mismo debe buscar articular es una concepción de persona que permita afirmar de modo categórico que la justicia tiene primacía. En segundo lugar, el argumento se mueve desde este entendimiento del *equilibrio reflexivo* a una interpretación de los elementos de la *posición original*. Si esta es una herramienta para obtener el *equilibrio reflexivo* allí deben encontrarse modelados los elementos que buscan equilibrarse. Puesto que uno de estos elementos es una concepción de persona de

<sup>21</sup> Afirma Sandel: “In the assumption of mutual disinterest, we find the key to Rawls’ conception of the subject, the picture of the way we must be to be subject for whom justice is primary. But the notion of the self as a subject of possession, taken alone, does not complete the picture. As the account of plurality suggests, no just any subject of possession will do, but only an antecedently individuated subject, the bounds of whose self are fixed prior to experience. To be a deontological self, I must be a subject whose identity is given independently of the things I have, independently, that is, of my interests and ends and my relations with other” (Sandel, 1982, p. 55).

la que pueda afirmarse de modo no contingente que otorgará *primacía a las exigencias de justicia* –una concepción metafísica de persona–, tal concepción debe encontrarse modelada en la *posición original*. Finalmente, el argumento se mueve desde esta interpretación de la *posición original* al problema de la *neutralidad*, para mostrar de qué modo este compromiso metafísico priva a la concepción de justicia rawlsiana de su carácter neutral. El movimiento del argumento va, pues, desde la *prioridad de lo correcto* al *equilibrio reflexivo*, y desde allí –pasando por la *posición original*– hasta la *neutralidad*.

Esta reconstrucción permite apreciar que el peso del argumento de Sandel recae por completo en su interpretación de la *prioridad de lo correcto en relación con lo bueno*. Si esta prioridad es entendida –tal como Sandel hace– como equivalente a la *primacía de la justicia*, entonces, como hemos mostrado, las conclusiones que este extrae se siguen. Por esta razón pensamos que, de las distintas réplicas que han sido dirigidas a Sandel, la formulada por Brian Barry es la que posee mayor relevancia. En efecto, Barry ha cuestionado la interpretación de la *prioridad de lo correcto* como equivalente a la *primacía de la justicia*. Ha señalado al respecto:

Michael Sandel's critique in *Liberalism and the Limits of Justice* trades attributing to Rawls an idea –“the primacy of justice”– that he does not employ and treating it as a portmanteau of the “priority of right” and “congruence” (Barry, 1995, p. 883).

Barry ha mostrado cómo Sandel erróneamente interpreta la *prioridad de lo correcto* como si tuviese implicancias motivacionales, como si esta tesis señalase que las exigencias de justicia tienen más peso –entendido como mayor importancia moral y mayor fuerza motivacional– que otros intereses morales y políticos. La tesis de la *prioridad de lo correcto*, sostiene Barry, tiene que ver con la manera de construir la concepción de justicia y no implica ninguna toma de posición con respecto a ninguna cuestión motivacional. Se refiere a cuál es el modo adecuado de elucidar los criterios que van a determinar qué exigencias debe satisfacer nuestro trato con los demás, no se refiere ni a la importancia moral de estas exigencias ni a su relevancia motivacional en comparación con otras consideraciones morales. Lo único que se limita a afirmar esta tesis es que la forma correcta de tratar a los demás no puede ser determinada a partir de la maximización de un estado de cosas que se considera va-

lioso con independencia de cualquier restricción moral o consideración por los otros.

Teniendo en mente las observaciones de Barry, es posible ahora coordinar las réplicas a Sandel formuladas por Freeman y Rawls.

En primer lugar, dado que la función del *equilibrio reflexivo* es justificar la concepción de justicia –y dado que no es una de las tesis de esta concepción que las exigencias de justicia tengan un peso motivacional inderrotable– no es necesario suponer que lo que busca equilibrarse es una concepción moral de justicia con una concepción antropológica de hombre para el cual aquellas exigencias deban tener tal carácter de inderrotabilidad. La concepción de justicia que busca ser justificada por el recurso del *equilibrio reflexivo* es una concepción de tipo moral que no contiene tesis en relación con la motivación humana. Estas consideraciones despejan el camino para entender el recurso del *equilibrio reflexivo* tal como lo hace Rawls. El objetivo de esta herramienta, señala, es articular en una concepción coherente las distintas convicciones que tenemos con respecto a la justicia. Es decir, no se buscan articular juicios morales con juicios antropológicos –como parece sostener Sandel– sino solo convicciones morales, de distinto grado de generalidad, referidas a un tema específico: la justicia de las instituciones sociales básicas. Afirma Rawls refiriéndose al método en general:

People have considered judgments at all levels of generality, from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions. One tries to see how people would fit their various convictions into one coherent scheme, each considered conviction whatever its level having a certain initial credibility (Rawls, 1975, p. 289).

En segundo lugar, es posible también comprender qué ha llevado a Sandel a cometer el error denunciado por Freeman. Su crítica específica es que Sandel:

ignore(s) the account of the person Rawls gives in *A Theory of Justice*, and substitute for it the account of the parties in the Original Position [...]. Sandel completely ignores the role of the moral powers in defining the person, as well as Rawls's claims regarding the connection of persons with their conception of the good (Freeman, 1994, p. 620, n. 8).

*A Theory of Justice*, señala Freeman, intenta dar respuesta a

tres problemas distintos: a) cuál es la concepción de justicia más apropiada para la estructura básica de una sociedad democrática, b) cuáles son las instituciones democráticas más adecuadas para satisfacer las exigencias de la concepción de justicia, y c) cómo explicar que las personas que crecen en una sociedad regulada por estas instituciones adquieran la voluntad o el deseo de actuar conforme a la justicia<sup>22</sup>. Sandel confunde, continúa diciendo Freeman, el problema al que intenta dar respuesta la primera parte de *A Theory of Justice* –para el cual se utiliza el recurso de la *posición original*– con el problema de la tercera parte de esta obra, el cual se refiere a si la concepción es plausible, esto es, si las personas que crecen en una sociedad ordenada por sus principios adquieren un sentido de la justicia normalmente efectivo.

Ahora bien, la razón de este error es que Sandel introduce *la primacía de la justicia* en sentido motivacional como un elemento de la concepción de justicia –que es objeto de la primera parte de *A Theory of Justice*– siendo que los problemas vinculados a la motivación, y las posiciones tomadas en este nivel, son tratados en la tercera parte de aquella obra. Es decir que la *primacía de la justicia* en sentido motivacional no forma parte de la concepción de justicia rawlsiana, sino que es algo que –luego de tener aquella concepción a mano– debe ser demostrado argumentando a favor de su estabilidad. Sandel confunde, como señala Barry, *la prioridad de lo correcto en relación con lo bueno* –prioridad que sí forma parte del argumento a favor de la concepción de justicia– con la *prioridad o primacía de la justicia* –tesis que Rawls nunca ha sostenido–.

Este artificio de entrecruzar la primera y la tercera parte de *A Theory of Justice* es el que le permite a Sandel, sostiene Freeman, encontrar corporizada en la *posición original* una concepción metafísica de persona, malinterpretando pasajes extraídos del argumen-

<sup>22</sup> Samuel Freeman propone un paralelismo entre estos tres problemas y los señalados por Immanuel Kant en su ensayo *Idea for a Universal History*. Señala: “Kant’s question raises then three kinds of problems, each suggesting a host of more specific issues. First, there is the problem of conception: we need an account and justification of the principles that define a just constitution. Second, we encounter an institutional problem (Kant’s problem of ‘experience’): how to describe the social and political institutions that are required to realize these principles in societies at particular periods. And third, a motivation problem arises: how to give an account of how people can acquire the will to do justice and the desire to support just institutions (as defined by answers to Kant’s first two problems)” (Freeman, 1994, p. 624).

to a favor de la estabilidad<sup>23</sup>.

En tercer lugar, y habiendo encontrado los motivos por los que la interpretación de la *posición original* ofrecida por Sandel es errada, queda despejado el camino para la interpretación ofrecida por Rawls.

Si la función de la *posición original* es la de una herramienta para alcanzar el *equilibrio reflexivo* y justificar así la concepción moral de justicia, entonces cada uno de sus elementos debe modelar alguna convicción en relación con la justicia. Esto garantiza que los principios generales obtenidos mediante este recurso estén en *equilibrio reflexivo* con aquellas convicciones. A su vez, como lo que se están buscando son principios de justicia para valorar instituciones que organicen la vida en sociedad, los juicios meditados se refieren tanto al carácter que deben tener las personas que viven en una sociedad justa, como a la configuración de esta última.

De esta manera, la *posición original* modela dos concepciones: una concepción de persona y una concepción de sociedad. Solo que a diferencia de lo sostenido por Sandel, y debido a que el objetivo del *equilibrio reflexivo* es articular juicios morales, estas concepciones no son metafísicas, esto es, no se refieren a lo que los sujetos o la sociedad *son*, sino normativas: se refieren a cómo los sujetos *deberían* ser tratados al momento de configurar las pautas que regirán su vida en sociedad<sup>24</sup>.

Finalmente, esta interpretación de las concepciones modela-

<sup>23</sup> Afirma Freeman: "It is, I believe, only when we reach chapter 9, and the second stage of the argument for stability, from 'congruence', that the deeper bases of the view in Kantian ethics become really apparent" (Freeman, 1994, p. 628).

<sup>24</sup> Es necesario aquí formular una aclaración. Si bien Rawls ha puesto énfasis en señalar el carácter normativo de la concepción de persona modelada en la posición original, a partir de la publicación de "Justice as Fairness: Political not Metaphysical", no consideramos que esta haya sido una corrección introducida a su teoría para enfrentar las críticas de Sandel. Que la concepción de persona tenía carácter normativo estaba claro ya desde *A Theory of Justice*, de allí que basten los recursos presentes en esa obra para enfrentar la objeción de falta de neutralidad metafísica. Lo que no se encontraba claro allí era el ámbito de aplicación de esta concepción normativa de persona. A solucionar esta falencia, y no a enfrentar las críticas de Sandel, han ido dirigidas las correcciones efectuadas por Rawls a su teoría desde 1985. Su esfuerzo ha sido dejar en claro algo que no lo estaba en *A Theory of Justice*: que el dominio al que se aplicaba la concepción normativa de persona es solo el político.

En consecuencia, "Justice as Fairness: Political not Metaphysical" expone dos posiciones. Una señala que la concepción de persona tiene carácter normativo y no metafísico. Esto era algo que había sido afirmado por Rawls con anterioridad a las

das en la *posición original* permite preservar la *neutralidad* de la concepción de justicia. La función de las *circunstancias de la justicia* y el *mutuo desinterés* no es modelar una visión de la sociedad según la cual sus miembros son prioritariamente plurales y se relacionan con sus fines y vínculos como un sujeto de posesión. Por el contrario, apelar a las *circunstancias de la justicia* en el diseño de la *posición original* simplemente garantiza que los principios morales que se obtengan –mediante el recurso constructivista de la situación hipotética de elección– sean principios de justicia, o sea, principios cuyo objeto es determinar la manera en que deben distribuirse las cargas y beneficios en una sociedad. Si no se incorporaran en la descripción de la *posición original* las *circunstancias de la justicia*, los principios morales obtenidos no serían adecuados para resolver los aspectos distributivos de la estructura básica de la sociedad, es decir no serían principios de justicia en absoluto<sup>25</sup>.

Como señala Rawls, la *posición original* modela entonces, por un lado, las condiciones en las cuales se presentan los problemas que los principios de justicia están destinados a resolver, y por otro, las restricciones que en función de nuestras convicciones o juicios meditados consideramos razonables imponer en la elección de aquellos principios<sup>26</sup>.

Vistas de esta manera, las circunstancias de la justicia tienen un fundamento empírico, a saber, las condiciones históricas y sociales de una sociedad moderna y democrática donde pueden presentarse problemas para distribuir cargas y beneficios de la cooperación social. En *A Theory of Justice* este punto no está claro, puesto que allí parecen ser presentadas como las circunstancias de cualquier sociedad humana. Afirma Rawls: "But a human society is characterized by the circumstances of justice" (Rawls, 1971, p. 112).

No obstante, del objetivo general de esta obra puede inferirse

---

críticas de Sandel, y basta para enfrentarlas con éxito. Otra señala que la concepción normativa tiene carácter político. Esta es una tesis novedosa que viene a subsanar las falencias de *A Theory of Justice*, falencias que nada tienen que ver con las apuntadas por Sandel.

<sup>25</sup> Sostiene Rawls en relación con la idea de una concepción de justicia: "is to be regarded as providing in the first instance a standard whereby the distributive aspects of the basic structure of society are to be assessed" (Rawls, 1971, p. 8).

<sup>26</sup> Suscribiendo esta interpretación señala Rawls: "Now, as we have seen, the relations of mutual self-interest between the parties who are similarly circumstanced mirror the condition under which questions of justice arise, and the procedure by which the principles of judgment are proposed and acknowledged reflects the constraints of having a morality" (Rawls, 1971a, p. 206).

que la caracterización debe ser entendida como circunscripta a sociedades democráticas. Así, en el prefacio se sostiene:

Of the traditional views, it is this conception, I believe, which best approximates our considered judgments of Justice and constitutes the most appropriate moral basis for a *democratic society* (Rawls, 1971, p. XVIII) (las cursivas son nuestras).

Es decir, aunque *A Theory of Justice* es ambigua en este sentido, existen bases para la interpretación de las circunstancias de la justicia sostenida en obras posteriores por Rawls<sup>27</sup>.

Lo importante aquí es que, sea en relación con las sociedades humanas en general o solo con las sociedades democráticas modernas, las circunstancias de la justicia tienen por finalidad modelar los hechos que las caracterizan y que hacen posible y necesaria la cooperación, y, por ende, la justicia.

Como se recordará, Sandel descarta esta interpretación, pero las razones por las que lo hace son nuevamente la confusión denunciada por Barry entre la *prioridad de lo correcto sobre lo bueno*, de tipo justificatoria, y la *prioridad de la justicia*, de tipo motivacional. Para explicar su posición al respecto, Sandel sostiene:

But an empiricist understanding of the original position seems deeply at odds with deontological claims. For if justice depends for its virtue on certain empirical preconditions, it is unclear how its priority could unconditionally be affirmed. Rawls says that he borrows his account of the circumstances of justice from Hume. But Hume's circumstances cannot support *the priority of right* in the deontological sense. They are after all empirical conditions. To establish *the primacy of justice* in the categorical sense Rawls' claim requires, he would have to show not only that the circumstances of justice prevail in all societies, but that they prevail to such an extent that the virtue of justice is always more fully or extensively engaged than any other virtue (las cursivas son nuestras) (Sandel, 1982, p. 30).

El argumento de Sandel para descartar esta interpretación de la *posición original*, incluidas las *circunstancias de la justicia*, sería el siguiente: si el objetivo del argumento rawlsiano es obtener principios de justicia que tengan una primacía categórica, en sentido

<sup>27</sup> Para ver la interpretación definitiva ofrecida por Rawls en relación con la función de las circunstancias de la justicia cf. Rawls, 2001, p. 197.



motivacional, entonces en las premisas del argumento que conduce a ellos no deben tener cabida hechos contingentes. Esto porque los principios obtenidos de esta manera serían hipotéticos –solo tendrían valor motivacional si estas circunstancias fácticas se dieran– antes que categóricos.

Sin embargo, una vez que se advierte, como señalan Barry y Freeman, que el argumento confunde dos partes distintas de la obra de Rawls –aquella que tiene por eje la *posición original* y las *circunstancias de la justicia*, cuyo objetivo es obtener principios de justicia que estén justificados con independencia de cualquier concepción particular del bien (lo que se encuentra vinculado con la *prioridad de lo correcto*), y la que tiene por objeto mostrar que en una sociedad bien ordenada los ciudadanos tienen un *sentido de la justicia* suficientemente fuerte como para prevalecer en general sobre otras motivaciones–, la objeción a la “interpretación empirista” se desvanece.

Que la caracterización de las *circunstancias de la justicia* modele hechos referidos a la sociedad no vicia ni la *prioridad justificatoria de lo correcto en relación con lo bueno* –única prioridad que esta parte de la argumentación reclama– ni dice nada, a favor o en contra, de la estabilidad de la concepción –objetivo de la tercera parte de *A Theory of Justice*–.

No existe dificultad alguna en que la argumentación que conduce a los principios de justicia se valga de hechos. Esto es perfectamente compatible con la manera de Rawls de entender la teoría moral y política. Así sostiene:

Moral Theory must be free to use contingent assumptions and general facts as it pleases (Rawls, 1971, p. 44).

A modo de síntesis, puede señalarse que las circunstancias de la justicia no modelan aquellas características que el sujeto debe poseer para que la justicia tenga en él *primacía motivacional*, sino aquellas características que debe tener cualquier sociedad humana –o cualquier sociedad democrática moderna– en la cual la cooperación es a la vez posible y necesaria, y en donde se presentan problemas vinculados con la distribución de cargas y beneficios.

Por último, en relación con el supuesto del *mutuo desinterés* –al igual que en el caso anterior– lo que este modela en la *posición original* son las circunstancias en que se suscitan las cuestiones de justicia. Ahora bien, para que estas circunstancias se den no es necesario que los individuos tengan únicamente motivos egoístas, ni

que estén constituidos metafísicamente como individualizados con anterioridad a sus fines y vínculos. Lo único que el supuesto del *mutuo desinterés* modela es la idea de que cuando se efectúan demandas de justicia los ciudadanos se reconocen como representantes de intereses que merecen ser considerados. Al respecto señala Rawls:

The justice of practices does not arise until there are several different parties (whether we think of these as individuals, associations, or nations, and so on, makes no difference) who do press their claims on one another and who do regard themselves as representatives of interests which deserve to be considered. These conditions can obtain under the most varied circumstances and from any number of motives [...]. *The conjectural account involves, then, no particular theory of human motivation [...] simply incorporate into the conception of justice the relations between persons which set the stage for questions of justice* (las cursivas son nuestras) (Rawls, 1971a, pp. 205-206).

Es decir, uno puede sostener que las cuestiones de justicia no se presentan mientras que no existan distintas partes que formulen reclamos que pretendan merecer consideración, y no verse comprometido por esto ni con una teoría egoísta respecto de la motivación humana, ni con la concepción antropológica que sostiene que los individuos capaces de justicia tienen que estar individualizados con anterioridad a cualquiera de sus fines, relacionándose con estos como un sujeto de posesión.

Una vez esclarecida la confusión de Sandel entre *prioridad de lo correcto* y *primacía de la justicia*, su interpretación de que el supuesto del mutuo desinterés modela la relación que debe existir entre un sujeto –para quien la justicia tiene primacía– y sus fines –a los cuales considera objetos de posesión– pierde asidero.

## 6. Conclusión

Sandel ha cometido cuatro errores concatenados. En primer lugar, interpreta la *prioridad de lo correcto en relación con lo bueno* –de tipo justificatoria– como si equivaliera a la *primacía de la justicia* –de tipo motivacional–. En segundo lugar, esto lo conduce a sostener que *el equilibrio reflexivo*, como herramienta de justificación de la concepción de justicia, busca articular esta concepción de justicia con una concepción de persona que permita afirmar de modo categórico la *primacía de la justicia*. En tercer lugar, dado que la *po-*

*sición original* es una herramienta para obtener el *equilibrio reflexivo*, y dado que uno de los elementos que debe articularse en *equilibrio* es una concepción de persona que permita sostener de modo no contingente la *primacía a las exigencias de justicia* –esto es, una concepción de índole metafísica–, la misma debe encontrarse modelada en el diseño de la *posición original*. Por último, Sandel afirma que la concepción de justicia rawlsiana –por estar comprometida con tal concepción metafísica de persona– no es *neutral*.

Como esperamos haber mostrado, de las réplicas dirigidas contra el argumento de Sandel, es la de Brian Barry la que posee mayor relevancia. Si la *prioridad de lo correcto* no equivale a la *primacía de la justicia*, tal como sostiene Barry, entonces la función del *equilibrio reflexivo* no es articular juicios morales y antropológicos. Como afirma Rawls, la única función del *equilibrio reflexivo* sería articular nuestras convicciones morales en relación con la justicia. La *posición original*, en calidad de herramienta para alcanzar este equilibrio, modelaría entonces –tal como sostiene Rawls– concepciones morales y no metafísicas. Según afirma Freeman, si Sandel piensa lo contrario es por haber confundido –al haber identificado de manera indebida la *prioridad de lo correcto* con la *primacía de la justicia*– dos partes distintas del argumento rawlsiano: la que busca determinar cuál es la concepción de justicia más adecuada para la estructura básica de una sociedad y la que busca explicar de qué modo las personas que crecen en una sociedad regulada por estas instituciones adquieren la voluntad o el deseo de actuar conforme a la justicia. Finalmente, todas estas consideraciones despejan el camino para la interpretación de los elementos de la *posición original* ofrecida por Rawls, la cual garantiza el carácter *neutral*, a nivel metafísico, de la concepción de justicia.

## Bibliografía

- Barry, B. (1989), *Theories of Justice*, California, University of California Press.
- (Julio 1995), “Rawls and the Search for Stability”, *Ethics*, 105, pp. 874-915.
- Daniels, N. (2000), “Reflexive Equilibrium and Justice as Political”, en: Victoria Davion and Clark Wolf (eds.), *The Idea of a Political Liberalism. Essays on Rawls*, Rowan & Littlefield Publishers Inc., pp. 127-154.
- Dworkin, R. (1975), “The Original Position”, en: Norman Daniels

- (ed.), *Reading Rawls*, Oxford, Basil Blackwell, pp. 16-53. Recogido también en R. Dworkin, (1977), *Taking Rights Seriously*, Cambridge, Harvard University Press, pp. 150-183.
- Freeman, S. (1994), "Political Liberalism and the Possibility of a Just Democratic Constitution", *Chicago-Kent Law Review*, vol. 69, N° 3, pp. 619-668.
- Lyons, D. (1975), "Nature and Soundness of the Contract and Coherence Arguments", en: Norman Daniels (ed), *Reading Rawls*, Oxford, Basil Blackwell, pp. 141-167.
- Rawls, J. (1971a), "Justice as Reciprocity", en: Samuel Gorowitz (ed), *John Stuart Mill: Utilitarianism, with Critical Essays*, Indianapolis, Bobbs-Merrill. Citado en la versión de *Collected Papers*, Samuel Freeman (ed.), Cambridge, Harvard University Press, 1999, pp. 190-224.
- (1971), *A Theory of Justice*, Cambridge, Mass., Harvard University Press. Edición revisada 1999.
- (1975), "The Independence of Moral Theory", *Proceedings and Addresses of the American Philosophical Association*, 48, pp. 5-22. Citado en la versión de *Collected Papers*, Samuel Freeman (ed.), Cambridge, Harvard University Press, 1999, pp. 286-302.
- (1985), "Justice as Fairness: Political not Metaphysical", *Philosophy and Public Affairs*, 14, pp. 223-252.
- (2001), *Justice as Fairness: A Restatement*, Cambridge, Mass., The Belknap Press of Harvard University Press.
- Sandel, M. (1982), *Liberalism and the Limits of Justice*, Cambridge, Cambridge University Press, segunda edición 1998.
- Seleme, H. (2004), "Coherencia, contractualismo y justicia procesal pura", *Revista de la Facultad de Derecho, UNAM*, T. LIII, N° 241.