Rawls, John (1921–2002)
George Klosko

John Rawls was a major twentieth century moral and political philosopher. His monumental *A Theory of Justice* (1971) provides a powerful defense of the liberal welfare state and the priority of individual liberty over other values. In *Theory* Rawls claims that justice is the “first virtue of social institutions” (1971: 3), and attempts to establish principles of distributive justice to govern the workings of society’s “basic structure,” its major political, social, and economic institutions. These are the proper focus of distributive justice, because of the decisive influence they exert on people’s life prospects. *Theory* is credited with reviving normative political philosophy among Anglo-American philosophers and, since its publication, has served as a model for the practice of political philosophy. But in spite of the work’s enormous influence, Rawls became dissatisfied with central aspects of its argument. In a series of articles culminating in *Political Liberalism* (1993), he worked out an alternative method to establish his central conclusions.

Rawls was born in Baltimore to a prominent family (for biographical details, see Pogge 2007: ch. 1). His early interest in religion and plan of possibly entering the priesthood changed as a result of his experiences in World War II, including encountering the evils of the Holocaust, which shook his faith (Rawls 2009: 261–3). He received his PhD in philosophy from Princeton University, and taught at Cornell University and MIT, before joining the faculty of Harvard University, in 1962, where he taught for more than thirty years.

**Rawls’s Contract Argument in “Justice as Fairness”**

A major reason for the enormous attention *Theory* received is Rawls’s revival of the social contract, as presented in the works of Locke, Rousseau, and Kant. Rawls claims that proper principles of justice are those that would be accepted by participants under the fair conditions of a hypothetical choice situation. Because the principles emerge from fair conditions, Rawls calls his theory “justice as fairness.” In traditional social contract arguments, such as those of Hobbes and Locke, individuals are imagined in a pre-political state of nature, which they move out of by choosing institutions of government. In such constructions, theorists are often unclear about whether the state of nature is intended as historically accurate or purely as a hypothetical choice situation.
Moreover, as readers of Hobbes and Locke know, by changing conditions in the state of nature, one can alter the form of government that will be chosen. One must therefore defend the plausibility of the conditions under which the choice is made as well as what is chosen. Rawls departs from the traditional social contract framework in arguing that the objects of agreement are principles of justice rather than political institutions. And, like Kant (Kant 1970: 79), he posits a purely hypothetical situation. Part of Rawls’s task is to devise a plausible choice situation, so that selection of principles will be convincing.

Rawls first employed a contract argument in his 1958 article “Justice as Fairness” (1958). But the project of devising a plausible choice situation for moral principles can be traced back to his PhD dissertation, which was the basis for his first published article, “Outline of a Decision Procedure for Ethics” (1951). In this piece, Rawls describes the qualities that must be possessed by an appropriate judge for making moral decisions. These are to a large extent obvious. The judge should be intelligent, well informed, have adequate time to make his decision, be sane, as opposed to suffering from mental illness, and not have personal interests in the case. Decisions made by judges with these qualities will have greater claims to validity than those by judges who lack them. Although Rawls does not argue that decisions by appropriate judges are true, they are clearly more likely to be true than decisions by defective judges.

Rawls’s attempt to devise a plausible choice situation in “Justice as Fairness” requires an account of the individuals who will choose the principles and the conditions under which they choose. The approach used here presents features central throughout Rawls’s career. Appropriate principles are those selected by choosers who are motivated by rational self-interest (more exactly, to promote the interests of the parties they represent). The moral quality of the principles is provided by the requirement that choice of principles be unanimous, agreed upon by all choosers. This approach is in keeping with Rawls’s view of justice as restraints imposed on individuals’ pursuit of self-interest for the sake of other people (for difficulties inherent in this approach, see Barry 1989).

It is important to recognize that, in describing his choosers, Rawls does not attempt to depict actual people; he does not provide an account of human nature. Rather, as in “Outline of a Decision Procedure,” the persons choosing the principles of justice have the qualities we recognize as appropriate for people assigned this task. To some extent these qualities too are obvious. The choosers should be rational in the economic sense, that is, wishing to maximize their interests. They should know their own interests, have a large degree of common interests, which is necessary for living in society together, and be roughly equal in power, so no one can impose his or her preferred principles on others. Perhaps more controversially, they should also not be envious. If the choosers are envious, the eventual principles they agree upon must entail complete equality, as any departure from equality will be unacceptable to those who are worse off. In making the choosers non-envious, Rawls’s reasoning is based on the possibility that all participants may benefit by allowing departures from equality. As long as everyone benefits, pure equality is not the most desirable outcome.

Rawls believes that two principles of justice will be chosen. The first guarantees equal and maximum liberty to everyone. The second has two parts. Economic arrangements should be beneficial to everyone, and all people should have equal opportunities to achieve different positions. While the reasoning for the first principle seems relatively non-controversial – though we will see problems that were eventually pointed out – the second is widely viewed as less compelling. We will return to this subject below.

In addition to the fact that the principles are chosen in a fair situation, Rawls believes they have the important attribute of transparency. Because the principles are fair to all members of society, they can be justified to all. Members of society “can face one another openly and support their respective positions, should they
appear questionable, by reference to principles
which it is reasonable to expect each to accept”
(1958: 178). This transparency distinguishes
liberal societies from others in which inequal-
ities are not justified at all, or perhaps attrib-
uted to forces beyond human control (Waldron
1987). Accordingly, Rawls’s conception of
fairness is closely akin to hypothetical consent.
The principles of justice are those inhabitants
of society would choose if they had the oppor-
tunity to do so: “It is this notion of the possi-
bility of mutual acknowledgment of principle
by free persons who have no authority over one
another which makes the concept of fairness
fundamental to justice” (Rawls 1958: 179).

Rawls’s position could just as easily be called
“justice as mutual acknowledgment.”

From “Justice as Fairness” to Theory of
Justice

In spite of its impressive achievements, the
argument of “Justice as Fairness” is seriously
flawed. The main problem is that, motivated
by rational self-interest, choosers will attempt
to craft principles that are beneficial to them-
selves and do not have incentives not to do so.
For instance, if you know that you are unusually
talented, you will insist on principles that
disproportionately reward talented people. If
you know that you lack talent, you will
favor principles of the opposite sort. If all
participants reason in this manner, the only
way agreement can be secured is if those
with greater bargaining power impose their
preferences on other people. Under those
conditions, the principles eventually agreed
upon would not be true principles of justice
but would reflect differences of bargaining
power in the choice situation. Rawls clearly
came to recognize this problem. As he writes
in Theory, “to each according to his threat
advantage is not a principle of justice” (1971:
141). The movement from the relatively simple
argument of “Justice as Fairness” to the massive
and massively complex theoretical structure of
Theory of Justice is in large part to get around
this difficulty.

To overcome this problem, Rawls introduces
into the choice situation – now famously called
the “original position” – one of his most cele-
brated innovations. He places the representa-
-tive individuals who select moral principles
behind a “veil of ignorance.” They are deprived
of knowledge about their particular character-
istics: for example, age, race, religion, talents,
where they are from, even the generation to
which they belong. Unacceptably self-interested
choice of principles is prevented by people not
knowing their own interests. With the veil of
ignorance in place, “there follows the very
important consequence that the parties have
no basis for bargaining in the usual sense”
(139). In addition, the parties in the original
position do not know their own conceptions
of the good. The apparent paradox that principles
of justice result from choice based on self-
interest rather than moral motivation is
explained by the need for the parties to reach
an agreement. Given inevitable disagreements
over conceptions of the good, if they were
motivated to pursue their moral ideals,
agreement would not be possible.

However, introduction of the veil of
ignorance raises a difficulty of another sort.
The participants in the original position are
supposed to choose principles that maximize
their interests. But how can they do this if they
do not know what their interests are? Rawls’s
solution is to introduce the concept of “pri-
mary goods,” “things that any rational man is
presumed to want,” because they are necessary
Among primary goods, Rawls includes lib-
erties, income and wealth, opportunities, and
the social bases of self-respect. Under the
assumption that more primary goods are pref-
erable to less, the process of choice becomes
selecting principles that provide the largest
possible package of primary goods. In addition,
although the representative individuals do not
know which society they are from, they are
allowed to know how societies in general
work. They are given knowledge of “the
general facts about human society,” including
theories in economics, politics, and other fields
(1971: 137–8). We should note Rawls’s ready assumption that there are non-controversial theories in these areas. At any rate, with these elements added, the choice situation has been radically transformed.

Critique of Utilitarianism

In “Justice as Fairness” and other important works, one of Rawls’s main concerns is to develop an alternative to utilitarianism, which was the then dominant approach in moral and political philosophy. Rawls’s main argument against utilitarianism is that not it but justice as fairness would be preferred in the choice situation. Utilitarianism would be objected to by those people who would lose out under resulting social arrangements. The fact that other people would be better off would not motivate them to sacrifice their own interests.

In *Theory*, Rawls commends utilitarianism for philosophical rigor. However, he believes that utilitarians often provide wrong answers to normative questions. Rawls believes their main problem is failure to recognize “the distinction between persons” (1971: 27). Utilitarians’ method of weighing pleasures and pains is appropriate for decisions by a single individual. For instance, although I find going to the dentist unpleasant, I may decide to go. Long-term benefits of avoiding tooth decay and gum disease far outweigh temporary discomfort. But if similar reasoning is applied to society, the benefits some people receive may come at a cost to other people, who do not themselves benefit.

The approach to moral questions that Rawls calls “intuitionism” does not rely on intuition in a strong sense, that is, a special faculty that perceives aspects of moral situations not perceptible by the five senses. Rather, he uses the term in reference to “an irreducible family of first principles” that we recognize as true and have to be weighed against one another (1971: 34). Common examples include principles according to which one should tell the truth, keep one’s promises, and not harm other people. But intuitionism lacks a means to weigh such moral principles against one another when they conflict. Attempting to combine the strengths of these two approaches, Rawls reworks intuitionism to make it more rigorous, while still preserving its ability to provide answers that are intuitively correct. In contrast to intuitionism’s inability to establish priorities among moral rules, Rawls’s argument is intended to establish the priority of liberty to other values.

The Principles of Justice

Rawls argues that, under conditions of uncertainty behind the veil of ignorance, the representative individuals will choose conservatively. Aware that, after the veil of ignorance is lifted, they may find themselves in unfavorable positions in society, they should pursue a “maximin” strategy, according to which the position of the worst off members of society is as advantageous as possible. The two principles that are chosen are slight modifications of those selected in “Justice as Fairness.” Once again, the first principle provides equal and maximum liberty for all members of society. The second is in two parts. The first part guarantees equal opportunity; the second ensures that economic goods are distributed in a way that is of maximum benefit to “the least advantaged” members of society. This is Rawls’s famous “difference principle.” Movement from general advantage, as presented in the second principle in “Justice as Fairness” to this focus on the least advantage is a notable change in Rawls’s position, although he believes it will make little practical difference. Conditions that are of maximum benefit to the least advantaged should be similarly beneficial to other members of society (1971: 80–2). In addition, Rawls argues that the first principle takes “lexical priority” over the second, namely, that it is not permissible to trade off liberty for other values.

The implications of the difference principle are central to Rawls’s justification of the welfare state. According to this principle, if we envision two hypothetical societies, A and B, even if A has a larger economy than B and so higher
incomes for most members of society, if the position of the least advantaged in B compares favorably to that in A, B is preferable. Rawls does not take a position on whether the property system should be capitalistic or socialistic. This question, like others concerned with the design of institutions, should be settled on the basis of conditions in society, which are revealed in stages, as the veil of ignorance is lifted. However, given emphasis on equality and the needs of the least advantaged, the principles of justice imply societies along the lines of Scandinavian social democracies. This is in contrast to a society that is more oriented toward the free market with a smaller social welfare component and so greater economic inequality, such as the United States.

One reason Rawls views the difference principle as just is his belief that the distribution of natural talents is morally arbitrary. According to Rawls, people do not deserve the attributes and abilities that allow them to be successful. These traits are largely accidents of birth, as is also the accident of being born into a family that will nurture and help one develop them. Even if a person succeeds by working hard, she does not deserve the traits of character that allow her to behave in this way. It follows that Rawls views people's talents and abilities as to some extent a social resource, to be used by society as a whole for the benefit of those who lack such talents. This controversial position is the basis for an important criticism of Rawls, that his theory lacks a proper concept of desert.

**Criticisms of A Theory of Justice**

Rawls's arguments for his two principles of justice generated enormous criticism. Major contributions by some of the most important political philosophers of the second half of the twentieth century are largely in the form of criticisms of Rawls. To cite four examples, Ronald Dworkin contends that Rawls's view does not sufficiently recognize differences between deserved and undeserved equality. Dworkin attempts to remedy this defect in a brilliant alternative proposal (1981a, 1981b).

G. A. Cohen criticized Rawls's focus on basic structures. Arguing for a view of distributive justice more egalitarian than that of Rawls, Cohen views the central focus of a just society as just individuals (Cohen 1989, 1997, 2009). Amartya Sen criticizes Rawls's claim that distributive justice should focus on distribution of resources. Sen contends, instead, that justice should address uses to which resources are put, people's equal ability to develop important capabilities, for which different people require different resources (Sen 1980). Probably the most influential response to Rawls is the libertarian alternative presented by Robert Nozick, Rawls's colleague in the Harvard Philosophy Department. In *Anarchy, State, and Utopia* (Nozick 1974), Nozick develops a historically based theory of distributive justice. To distribute resources according to abstract moral principles is to overlook the fact that resources come into the world attached to the people who create them. Proper principles of justice should address the means through which people acquire resources in the first place and voluntary exchanges with other people.

H. L. A. Hart raised important objections to the treatment of liberty in Rawls's first principle of justice (Hart 1975). There are obvious problems with Rawls's claim that liberty cannot be traded off for other values. In fact, liberty must be regularly traded for other values in any system of private property. For example, your property rights to your backyard generally preclude my pitching a tent and camping in it. Moreover, Hart believes that Rawls fails to provide a means of reconciling conflicts between liberties themselves. For this purpose, Rawls requires a fuller account of the person, which would guide us in determining which liberties are more important than others.

In regard to Rawls's second principle, a widespread criticism is that the difference principle would not be chosen behind the veil of ignorance. According to standard probabilistic reasoning, if people lack information about their situations, they should choose to maximize average shares. Rawls's main consideration in support of his choice is that
other alternatives are overly risky (esp. 1971: 154–5) But theorists have responded with a mixed principle, which combines an adequate minimum level (poverty line) below which people cannot sink and maximizes average share above that level. Because this position avoids the extreme risk that is central to Rawls’s rejection of a pure utilitarian position, critics believe it would be chosen over the difference principle in the original position (Barry 1973: ch. 8, 1989: ch. 6; Harsanyi 1975; Fishkin 1975). It should be noted that, in defending his principles, Rawls appeals to considerations beyond the strictly economic. A perhaps under-recognized feature of his thought is concern that moral principles be able to generate their own support – in his terms, that they be “stable.” One reason why the difference principle is preferred to possible alternatives – those considered in Theory, which do not include a mixed principle – is the representative individuals’ recognition that they will be superior in this respect. But Rawls does not make clear how the representative individuals should weigh the advantages of the two principles in this respect against the relatively larger size of the packages of primary goods that would result from other principles.

Although Rawls’s account of choice in the original position is not based on the psychology of actual people, it is interesting to note experiments with actual people placed behind a veil of ignorance, conducted in different countries – the US, Canada, and Poland. Participants overwhelmingly chose a mixed principle, 77.8 percent of the time. A principle similar to the difference principle was chosen only 1.2 percent of the time (Frohlich and Oppenheimer 1992, p. 60).

An additional criticism developed by Rawls himself concerned the “stability” of justice as fairness. As just noted, in a “well-ordered society,” moral principles are stable. People internalize and behave in accordance with them. However, at some point, Rawls came to believe that the argument of Theory does not adequately recognize the pluralism of liberal societies. Contemporary liberal societies are characterized by inhabitants who belong to widely different cultural and religious traditions and hold different philosophical views – what Rawls calls “comprehensive” views. In Theory, the principles of justice are defended from a particular perspective and so are in effect supported by a particular comprehensive view. At some point, as we will see, the need to develop principles that could be accepted by proponents of different comprehensive views led Rawls to seek a dramatically revised approach to justice as fairness.

**Reflective Equilibrium**

In the middle of the twentieth century, political theory in the Anglo-American tradition was widely viewed as “dead,” as it was famously described by Peter Laslett. According to Laslett, there was a glaring absence of major works, in comparison to previous periods (1956: vii). But according to many scholars, this situation changed in 1971, with the appearance of Rawls’s “Olympian” work (Laslett and Fishkin 1979: 1).

Rawls’s revival of political philosophy was due as much to his methodological advances as to his substantive contributions. The contract argument in Theory is actually subordinate to a more basic method that can be traced back to Rawls’s doctoral dissertation and “Outline of a Decision Procedure for Ethics” (1951). As we have seen, in the latter, Rawls is concerned with the perspective from which to make proper moral judgments. In Theory, Rawls calls judgments that have the necessary attributes “considered judgments.” Moral beliefs in which we have greatest confidence should serve as “provisional fixed points” around which to construct overall moral theories. Examples Rawls provides in Theory are that racial discrimination and religious intolerance are wrong. If an account of justice is inconsistent with these convictions, it is not likely to be correct (1971: 19–20).

Rawls calls his method of moral justification “reflective equilibrium.” As a method of justification, it is concerned with establishing
coherence within one's moral beliefs. Reflective equilibrium is a practical method and does not address controversial metaethical issues such as the nature of ethical judgments and whether moral truth is possible or accessible. Rather, reflective equilibrium attempts to create an integrated structure of moral beliefs and principles. General moral principles should imply results that are intuitively sound, while one's understanding of particular cases should be consistent with intuitively sound principles. In order to achieve consistency, one must work back and forth between principles and particular cases, revising one's understanding of each to achieve overall coherence. Thus Rawls describes moral theory as “Socratic” (1971: 49). He does not believe it is possible to deduce a conception of justice from self-evident first principles. Rather, “its justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view” (1971: 21).

As noted above, Rawls is aware of the possibly arbitrary nature of contract arguments. By changing the features of the original position, one is able to change the principles that will be chosen in it (1971: 141). Thus, as noted, in order to be successful, Rawls’s contract argument must not only demonstrate that his two principles will be chosen, but it must also establish the plausibility of his account of the original position. By adjusting the account of the original position so that it will lead to selection of principles of justice that are also intuitively plausible, one may establish an equilibrium between the contract argument and our intuitive judgments. The equilibrium is “reflective” because, in order to attain it, we must critically examine all components of the argument, revising them as appropriate. Because his argument as a whole is subjected to a test of overall intuitive plausibility, this takes precedence over the contract argument. As his philosophical career continued, Rawls’s different versions of the contract argument received less attention and overall plausibility more.

As an approach to ethical and political problems, reflective equilibrium is not precisely defined. Its adherents disagree about specific features, and do not always claim to follow Rawls. However, regardless of such disagreements, since the publication of Theory, a great many theorists have practiced a similar style of political philosophy – generally referred to as “analytical,” “liberal,” or “Rawlsian” theory (for good brief discussion, see Miller and Dagger 2003). Adherents of this method prize ideals of rigor and clarity. Terms they use are carefully explained and questions of moral foundations generally avoided. These theorists argue from considered convictions they believe to be shared by their readers, making use of carefully crafted examples to sharpen intuitions. Practitioners of this style of argument address practical issues, and one of Rawls’s great accomplishments was to provide a model for how practical normative issues could be approached.

Like other aspects of Theory, the method of reflective equilibrium has received extensive criticism. One objection is that the intuitions from which practitioners proceed are products of society, not in any way basic. Thus the conclusions the method produces reflect no more than beliefs inculcated by society. Rawls anticipated this line of argument in “Outline of a Decision Procedure.” Reflecting the practical orientation of his theorizing, he claims it is obvious that the most plausible approach to moral questions is through ideas in which we have greatest confidence. It would be foolish to proceed from other ideas in which we have less faith (Rawls 1951: 196). The moral beliefs on which reflective equilibrium builds are not “arbitrary commitments or sentiments that we happen now to share,” to use one critic’s words (Lyons 1975: 146), but moral convictions in which we have greatest confidence. Moreover, it is likely that critics themselves share these beliefs. In the final analysis, reflective equilibrium is an ad hominem method, forcing the critic to admit that he too could not reject the points at issue, without conflicting with other moral beliefs he sincerely holds.

An additional line of criticism is perhaps more troubling. Is it true that the reader will
share Rawls’s moral intuitions? In *Theory*, he says that he does “not even ask” if one person’s sense of justice is the same as those of other people (Rawls 1971: 50). As we will see momentarily, in the years following publication of *Theory*, Rawls came to be increasingly troubled by the fact of disagreement and significantly revised justice as fairness to take it into account.

So great has been Rawls’s influence that *Theory of Justice* should be considered a “paradigm” for subsequent work in political philosophy, in the sense made popular in Thomas Kuhn’s famous work, *The Structure of Scientific Revolutions* (1962). Kuhn argues that significant scientific experiments fundamentally alter the way science is practiced. Confronted with the great works of a Copernicus or a Newton, subsequent researchers use it as a model for their own work, deriving from it their ideas of the questions important to examine, models for how to go about doing research, and what constitutes an adequate explanation. In all these respects and others, subsequent political philosophers have based their own practice on Rawls’s model. This in large part explains Rawls’s enormous influence and the central place his works occupy in contemporary political philosophy.

**Political Liberalism**

Among the many criticisms *Theory* received, Rawls found several particularly troubling and attempted to revise his theory to meet them. Several important lines of criticism have been mentioned above, especially Rawls’s failure to consider adequately the pluralism of contemporary liberal societies, which had damaging implications for the “stability” of justice as fairness (1993: xv–xviii). To address this problem, Rawls recast his position as a “political” liberalism. In his pioneering articles, “Justice as Fairness: Political Not Metaphysical” (1985) and “The Idea of an Overlapping Consensus” (1987), Rawls claims that justice as fairness is intended to serve a practical function, allowing people to live together cooperatively in pluralistic societies, rather than to be an expression of philosophical truth.

As Rawls uses the term, a “political” conception addresses a narrow range of practical issues. This is in contrast to a “comprehensive” view, which is an overall doctrine that provides answers to a full range of philosophical questions. Familiar comprehensive views are religious and other overall philosophies of life. Rawls believes that the pluralism of liberal societies is a natural result of unconstrained inquiry. When people deliberate about moral and political issues, they focus on different aspects of questions and approach them through diverse methods of investigation. Thus reasonable people come to different and irreconcilable answers (Rawls 1993, 54–8). Rawls labels this fact of human reason the “burdens of judgment.” Diversity of doctrines is “a permanent feature” of liberal public culture (Rawls 1993: 216–17). It can be prevented only through oppression, and so only if it is accepted can liberal society be legitimate. As in his earlier work, Rawls is deeply concerned with people being able to understand and accept the moral principles under which they live. He describes the “liberal principle of legitimacy” as follows:

> our exercise of political power is proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideas acceptable to them as reasonable and rational. This is the liberal principle of legitimacy. (Rawls 1993: 217)

This ideal can be realized and society function effectively only if people agree on a set of central moral principles – a “political” conception of justice. Because of their overall disagreements, each person will understand these principles from the perspective of his or her own comprehensive view. The result is what Rawls calls an “overlapping consensus.” He believes that justice as fairness is one set of principles that can fulfill this function. If justice as fairness is to be accepted by the diverse
inhabitants of a pluralistic society, it cannot be derived from one particular philosophical perspective – as it is in Theory – but must be worked up from elements of the public culture of liberal societies that are recognized by all reasonable inhabitants. The main “intuitive ideas” that fit this bill are conceptions of the person as free and equal and of society as a cooperative enterprise for mutual advantage. Choice of principles of justice in this newly conceived original position should reflect these basic ideas.

In addition to having this particular content, Rawls believes an overlapping consensus should be accepted by its adherents in a particular way. Rawls contrasts the overlapping consensus with a different form of agreement, a “modus vivendi.” This is conceived on the model of a truce or the outcome of political bargaining. Under a modus vivendi, citizens accept a set of principles for self-interested reasons. Its terms are determined by the balance of power between contending factions, and so may be renegotiated as the balance shifts (Rawls 1993, 147). Unlike a modus vivendi, an overlapping consensus is a principled agreement. Because its adherents accept it for moral reasons, they will not withdraw their support even if their views gain support and they can negotiate more favorable terms (Rawls 1993, 147–8). Because the balance of power on which a modus vivendi rests is liable to shift over time, Rawls believes the distinction between modus vivendi and overlapping consensus has important implications for social stability.

By the time he conceived Political Liberalism, Rawls had come to believe that an overlapping consensus can exist only in a liberal society. Only in such a society, does the majority of inhabitants accept a suitable political conception of justice and adhere to it in the proper way. Rawls’s view in this regard breaks with the apparently universalistic ambitions of Theory. Liberal societies have benefited from the long history of religious strife following upon the Protestant Reformation, which has led their inhabitants to recognize the need to set their religious differences aside and to cooperate with one another. Rawls believes that similar tolerance should be extended to philosophical differences. People should be willing to live cooperatively with others who differ from them in these respects. There are still members of liberal societies who are not sufficiently tolerant of disagreements and so not “reasonable,” in Rawls’s sense (1993: 48–64). While such people are included in liberal societies, their views need not be taken into consideration in deriving the moral and philosophical content of the overlapping consensus. Unless most inhabitants of a society subscribe to reasonable comprehensive views, political liberalism is not possible.

Central to the culture of liberal societies is a particular conception of the person as possessing two “moral powers”: ability to choose a conception of the good and to pursue it in a plan of life, and ability to live cooperatively with others people. Recognizing the two moral powers provides new direction to the representative individuals in a revised original position. The representative individuals are now viewed as motivated to promote development of the moral powers. The primary goods are reconceived as means for developing the moral powers, and so representative individuals are to choose principles that provide them with the largest possible share of those goods required for this purpose.

Rawls retains his belief that justice as fairness will be the result of the choice procedure. Recognition of the moral powers provides a response to Hart’s important criticism. In accordance with people’s overriding interest in developing their moral powers, different liberties can be indexed (1993: 291). For example, because freedom of conscience contributes to the moral powers more directly than do property rights, it takes precedence. In regard to the second principle of justice, Rawls to some extent loses faith in the difference principle. While he continues to believe it is the preferred principle of economic distribution, he recognizes that there is no knockdown argument for it in regard to other plausible principles and that the balance of considerations in its favor

In his late work, *Justice as Fairness: A Restatement* (2001), Rawls’s defense of the difference principle shifts to some degree. Rather than arguing directly that it appropriately maximizes shares of primary goods, Rawls focuses on how the difference principle contributes to a well-ordered society and the stability of liberal principles (2001: §34–40). But even if the difference principle is not necessarily the preferred principle of distribution in liberal societies, it is still one principle on which representative individuals can agree.

In order to be acceptable, Rawls believes that principles of justice must satisfy the requirements for liberal principles of justice. He views a liberal doctrine as comprised of three elements: a set of basic rights and liberties; assignment of special weight or priority to these rights as compared to other considerations, for example, overall considerations of social utility; means to ensure citizens have resources adequate to enable citizens to make use of their rights (1997: 774). The third item is central to the liberalism of the social welfare state, in contrast to earlier forms of liberalism and libertarianism. Although Rawls acknowledges different ways in which this condition can be satisfied, his arguments are intended to demonstrate that the difference principle fulfills these requirements.

A component of political liberalism that has attracted considerable attention is Rawls’s doctrine of “public reason.” This too is made necessary by the pluralism of liberal societies. Liberal legitimacy requires not only agreement on a political conception of justice but also that people accept commons standards of argument for discussing matters that involve the use of state power.

Principles of public reason too are chosen in the original position. In addition to selecting principles of justice, the representative individuals must agree upon the epistemological principles that are intended to define “public inquiry” (Rawls 2001: 89). Because these forms of argument must be acceptable to all reasonable citizens, public reason too must be “freestanding.” Its subject matter is confined to “constitutional essentials” and matters of basic justice. But when these matters are discussed, arguments used must be independent of particular comprehensive views. Acceptable forms of argument include formal logic, the sciences, and “presently accepted general beliefs and forms of reasoning found in common sense” (1993: 224, 1997: 773–80). As Rawls employs it, public reason serves mainly to limit the kinds of arguments that may be presented in public debate. It rules out forms of argument—for example, interpretations of religious texts—that are particular to specific comprehensive views and so could be expected to be acceptable to only certain inhabitants of pluralistic societies. An analogy to what Rawls has in mind is the rules of evidence used in courts of law. Even if some piece of evidence may be helpful in establishing the truth, it will be excluded if it does not meet criteria of acceptability, for example, if it was obtained illegally. And so one requirement for being a liberal citizen is ability to set aside one’s beliefs, perhaps one’s deepest beliefs, in order to address one’s fellow citizens only with arguments that they could be expected to accept.

Much of the criticism *Political Liberalism* has received has centered on public reason. His position is resisted especially by religious citizens, who believe that it is unfair to require that they set aside their deepest beliefs, while secular citizens are not required to do so. They view this as an infringement on their liberty (Wolterstorff 1997: 94; Eberle 2002: 232; Stout 2004: 68, 75–7), and claim that all citizens should be allowed to put forth their views, with differences to be settled by the political system (Wolterstorff 1997: 94; Eberle 2002: 232; Stout 2004: 68, 75–7). But Rawls fears the potential tyranny of the majority that could result.

**The Law of Peoples**

Rawls’s final major work, *The Law of Peoples*, addresses relations between rather than within societies. His task is to identify appropriate
principles of foreign policy that would be chosen by liberal societies and others that he describes as “decent.” According to Rawls, decent societies, while not liberal, protect the rights of their citizens, are directed by what he calls “common good” conceptions of justice, and, though not democratic, provide opportunities for general consultation. In this work, Rawls discusses “peoples” rather than nations or societies, although there is considerable overlap. Peoples, like these other entities, are described as sharing common cultures and moral sentiments and, like nations, control territories and have common institutions. But Rawls prefers the term “peoples,” because he believes the traditional notion of “sovereignty” is overly strong, as it allows states to go to war for self-interested reasons and to violate the rights of their citizens, regarding the latter as an internal matter (1999: 25–30).

To derive appropriate principles for the foreign policies of liberal societies, Rawls introduces a second original position, to be populated by representatives of liberal societies. Because questions concerning the proper principles of justice within these societies have been answered by the first original position, which has yielded Rawls’s two principles of justice, these matters are not addressed in this second original position. Confining attention to foreign relations, the representatives agree on familiar principles, including self-determination, non-intervention, the need to keep treaties, and the like. Rawls also includes a duty to assist other peoples that have “unfavorable conditions that prevent their having a just or decent political and social regime” (37). While the other principles agreed on are generally familiar norms of international relations, formal recognition of this duty of assistance would be a notable departure from current practice. In the second stage of this second original position, representatives of decent societies deliberate with each other and decide on the same principles. Thus Rawls concludes that the rules he identifies should regulate the international relations of liberal and decent peoples.

In addition to liberal and decent peoples, Rawls recognizes two classes of peoples that must be dealt with: “outlaw” peoples and “burdened societies.” The former are characterized by waging aggressive war or not respecting the human rights of their citizens. In order to promote compliance with international norms, liberal societies may take actions against outlaw societies, including the use of economic sanctions and going to war, in accordance with principles of just war, which Rawls discusses at length. Because of disadvantageous circumstances, burdened societies lack the ability to develop just or decent institutions. Rawls’s response to them is that liberal peoples should provide aid to address catastrophic immediate conditions and to help them develop the institutions necessary to provide acceptable lives to their citizens. In his discussions of both outlaw and burdened societies, Rawls’s aim is to help them eventually become “well-ordered” – full members of international society.

*Law of Peoples* has become a central text in current debates about global justice. As with Rawls’s other works, it has generated an enormous critical response. Much criticism has focused on the relatively undemanding redistributive principles Rawls defends. Critics have been puzzled by the way he frames his second original position. If, as Rawls says in *Theory of Justice*, justice entails an absence of arbitrary inequalities, then steps should be taken to remedy enormous disparities between living conditions and prospects between rich and poor countries, which are obviously arbitrary, and therefore unjust. Early readers of *Theory of Justice* imagined that a properly constructed original position would be global in extent. Behind the veil of ignorance, representative individuals would not know which societies they were from, and so resulting principles of distributive justice would apply between individuals in different societies rather than being confined to individuals within the borders of particular societies. Because of their strong desire to avoid the poor life prospects that would follow from being born into desperately poor countries, the representative individuals
would recommend a global difference principle (Barry 1973: ch. 12; Beitz 1979). Although Rawls briefly discusses international matters in *Theory*, he does so only in the context of discussing rights of conscientious refusal during times of war, and does not address questions of international distributive justice (1971: 377–9).

Rawls's reasons for his approach center upon a commitment to tolerance similar to that which motivated *Political Liberalism*. Because a global original position would result in principles of justice like those developed in *Theory*, imposing them on societies generally would be intolerant and not recognize the pluralism of international societies (1999: 59–62). In criticism of Rawls's position, it could be noted that this seems to run together matters of foreign policy – the advisability of sanctions and other measures – and normative questions concerning the principles that are just. According to the criteria used in his other works, decent societies, which do not treat people equally, are therefore unjust, and should be recognized as such, regardless of whether they should be openly condemned. Moreover, Rawls's attention to “decent” societies is itself puzzling, since there seem to be few of them in the world, in comparison to what he calls “benevolent absolutisms.” This category of societies is characterized as respecting human rights, though not affording their members significant opportunities for consultation (4, 63). Although these likely constitute a far more important class of societies, Rawls devotes virtually no attention to proper relations with them.

While the modest results of Rawls's inquiry in *Law of Peoples* may be explained by his intention to focus on the foreign policy of liberal societies, relations with burdened societies raise fundamental questions of distributive justice. Although Rawls's principles would impose additional redistributive duties upon existing liberal societies, many critics object to his claim that principles of justice hold only within societies, in contrast to less demanding duties of humanitarian assistance between societies. This distinction follows from Rawls's belief that inhabitants of individual countries are bound together by ties of justice, because of the cooperative relationships they enjoy, which are found only within societies. Moreover, Rawls contends that no society lacks sufficient resources or possibilities to become a full member of the society of peoples. Decisive is a society's political culture, and so how it uses its opportunities. For this reason, the duties of liberal societies toward burdened societies may be viewed as transitional, to help the latter develop appropriate institutions and accompanying culture to move out of the status of burdened societies.

Rawls's claims concerning the distinctiveness of cooperative relationships within societies is countered by theorists who contend that similar cooperative relationships exist in the international realm (Beitz 1979; Cohen and Sabel 2006). Scholars note the voluminous transactions that cross national borders. Examples include communications, travel, trade, foreign investment, and multinational corporations. A network of global institutions has also arisen, including the World Bank, World Trade Organization, and International Monetary Fund. According to these scholars, strong cooperation in the international realm analogous to that within states entails principles of economic justice in the international realm as well. Discussion in regard to these claims, as with multitudinous other aspects of Rawls's works are ongoing.

**Concluding Remarks**

In recent years, additional works of Rawls have been published. These include a brief, comprehensive and so compressed overview of his basic theory of justice, *Justice as Fairness: A Restatement* (2001), and volumes containing the lectures on the history of moral philosophy and political philosophy that he delivered to his classes at Harvard (2000, 2007). So great is interest in Rawls that his undergraduate thesis, *A Brief Inquiry into the Meaning of Sin and Faith*, has also been published, along with an essay on his religious views written toward the
end of his life (2009). As this body of material is considered by scholars, our understanding of different dimensions of Rawls's philosophy will doubtless increase, and it is likely that his claim to a central position in twentieth-century moral and political philosophy will be further strengthened.

SEE ALSO: Distributive Justice; Dworkin, Ronald (1931–2013); Equality; Hart, H. L. A. (1907–92); Justice; Liberal Theory; Liberalism; Liberty; Normative Theory; Nozick, Robert (1938–2002); Rights; Social Contract; State of Nature; Utilitarianism; Welfare State

References


Abstract

John Rawls was a major twentieth century moral and political philosopher. His monumental *A Theory of Justice* provides a powerful defense of the liberal welfare state and the priority of individual liberty over other values. In *Theory* Rawls claims that justice is the “first virtue of social institutions,” and attempts to establish principles of distributive justice to govern the workings of society’s “basic structure,” its major political, social, and economic institutions. These are the proper focus of distributive justice, because of the decisive influence they exert on people’s life prospects. *Theory* is credited with reviving normative political philosophy among Anglo-American philosophers and, since its publication, has served as a model for the practice of political philosophy.

Keywords: equality, equilibrium, justice, social contract