

John Rawls

Theory of Justice

By Avery Kolers, University of Louisville

Objectives

1. Explain why Rawls's theory of justice is first and foremost a *procedural* theory.
2. Present and explain the two principles of justice.
3. Identify Rawls's view of the relationship between individual and society, and his objections to rights-based, utilitarian, and communitarian views.
4. Address Rawls's omission of some of the most important issues of justice in contemporary human societies, such as race, colonialism, health, migration, and the global environment.

Reading Assignment:

Rawls, John. *A Theory of Justice*. Cambridge, Mass: Belknap/Harvard University Press, 1971.

Commentary

John Rawls's *Theory of Justice* (*TJ*) is the most important work of 20th century normative political philosophy, and "Justice as Fairness," the theory he defended there, is the most important normative theory in the field. A philosophical work is generally considered part of *political* philosophy when it addresses systems through which power is contested and exercised or resources are produced and distributed. A work of political philosophy is *normative* when it engages *moral* questions about these issues – for instance, how the exercise of power can be justified to those who are subject to it, or how resources such as money and power should be distributed, or whether it is ever permissible to trade off the lives and interests of some for the benefit of others. And finally, a work of normative political thought is generally considered political *philosophy* as opposed to political *thought* when it abstracts from the specific histories and institutional arrangements of particular locales, and attempts to isolate its normative political questions as they might apply in any society, anywhere. Thus a political philosopher might put the arguments of Aristotle, Hobbes, and Nussbaum into direct conversation as though there were neither time nor space separating them. It was this tradition of debate – *normative political philosophy* – that Rawls's *Theory of Justice* rejuvenated and reshaped upon its appearance in 1971.

Justification

During the 17th and 18th centuries, philosophers such as Thomas Hobbes, John Locke, and Jean-Jacques Rousseau adapted an older "Natural Law" tradition by using the image of a "social contract" to ask what

might justify political rule and property rights over land and goods. Their strategy was to describe their vision of a 'state of nature' – a world before government – and imagine what we would agree to if we found ourselves in such a world. For all its limitations, this tradition was revolutionary because it put governments on notice that in order for their rule to be legitimate, they had to provide benefits to the everyday citizen. Moreover, since *each* citizen was *equally* entitled to ask for such justification, the social contract tradition implied a basic moral equality of all, recasting unequal social relations as artificial rather than natural. Out of this tradition emerged theories of political and economic rights that are still relevant and compelling today, including the American Revolution's ideals of "life, liberty, and the pursuit of happiness," and the French Revolution's ideals of "liberty, equality, fraternity."

Even so, for over a century before *TJ*, the social contract tradition had waned, and Utilitarians and 'Intuitionists' dominated normative political philosophy. 'Intuitionists' are those whose arguments rely on direct appeals to moral considerations such as justice or equality (34). For instance, an intuitionistic defense of socialism might argue that equality of condition was more important than getting what you deserve; but a different intuitionist could equally well defend capitalism by arguing that individual property rights outweigh appeals to the social good. Rawls worried that these sorts of debates were futile because there was no external standard for determining which set of 'intuitions' about justice was correct (39). Utilitarians, by contrast, evaluate actions and social institutions by asking whether they maximize well-being for all affected. For instance, a utilitarian might endorse a policy of fast economic growth even at the cost of some people's basic interests in the short term. Rawls objected that utilitarianism is too dismissive of the distinction between persons (27). It's one thing for you to sacrifice a fun morning (by studying till noon) so you can have a fun evening; it's quite different when policymakers sacrifice *your* pleasure or even life so that *others* can have fun.

Against intuitionism, Rawls argued that progress in political philosophy could come, not from any great new insight about the *substance* of justice, but from devising a fair *procedure*. To Rawls, the shift to a procedural conception of justice brought numerous benefits, the first of which was that fair procedures help us make decisions when we can't agree on substantive outcomes. It is typically easier to see that an election, a lottery, or a coin-toss is fair than it is to determine which of the outcomes is better on its merits. Indeed, the justice of giving power to the winner of the election, or money to the winner of the lottery, is *purely* procedural: the outcomes are right *just because* they derive from a fair procedure (86). So in *TJ*, Rawls proposes, not to simply offer a theory of justice, but to describe a fair procedure for identifying or choosing principles of justice for the distribution of the benefits and burdens of social cooperation, and then to let the substantive chips of justice fall, more or less, where they may (120). (He does insist that the procedure doesn't work like a mathematical procedure or a logical deduction; whatever comes out of his procedure must still be double-checked from the perspective of those who are to live under the principles so chosen (499), as well as by reflecting on the system as a whole (579).)

Procedural justification has other benefits, too. First, a fair procedure makes it easier for everyone – losers included – to live with the results. We can reconcile ourselves to unfavorable outcomes by accepting that the procedure from which they arose was fair and legitimate. And second, by setting up, not a static society that strictly distributes a fixed amount to each person or class, but rather a set of economic and political procedures that play out across decades and generations (304), we can hope and expect that over time, each of us, or each of our family lines, might take turns in all strata of society.

The Veil of Ignorance and the Two Principles

What, then, was the procedure that Rawls thought we could all see would be fair? And what principles did he think would emerge from this procedure?

Traditional social contract theories are procedural theories that imagine everyone in a state of nature choosing a set of rules that would make it worthwhile for them to submit to governmental authority. Rawls has two reasons for rejecting the traditional method. First, although he agrees that each person is owed a justification for exercises of power over them, he thinks that we should not start from the false idea that we were once in a state of nature without government, or would go back there if there were no agreement. But second, any starting point *other* than this phony state of nature would lead us to skew the rules in our favor: if I am in the religious or racial majority, I can make the minority pay for tolerance; if I am independently wealthy, I can resist any demands for redistributive taxation.

To overcome this dilemma Rawls proposes that a fair procedure is one where each of us is hypothetically situated in an “Original Position” of equality where we know enough about people and societies to choose rationally about basic social systems, but know nothing particular about ourselves, our own society, or even our place in history (120). So what we *do* know are general scientific and social scientific facts about the nature of human beings and societies; we know that we live among humans in a situation of “moderate scarcity” – that is, we can all survive and thrive, but we have to work for it because we won’t just receive manna from heaven. And we know that people tend to want things like money, self-respect, social status, political and moral freedoms, and the opportunity to choose and pursue their own life plan. What we *don’t* know is anything particular about ourselves or the specific society in which we live: whether we are rich or poor, able to perform physical labor or give birth, a member of the majority ethnicity or of any religion, and so on. He calls this situation of not knowing particulars the “veil of ignorance,” and proposes that justice is whatever principles for governing our society we unanimously choose from behind this veil (136).

Because each of us can “enter” the Original Position at any time – we do not have to time-travel back to a past state of nature or move to an unsettled territory – we can ask ourselves the question of basic justice whenever that question arises. And since each of us is represented in the Original Position, we are not going to accept a sacrifice of our own interests just to benefit someone else. On the other hand, because we choose from under a veil of ignorance, everyone would reach the same conclusions about what justice is. If *everyone* chooses from under a veil of ignorance, then *each* of us has the same information and preferences, and *all* of us choose identically (139) and unanimity is achieved.

And what, then, do we unanimously choose? Rawls posits two principles of justice; the first has to do with what is to be held equal and guaranteed its “fair value”; the second has to do with what may be distributed unequally. For Rawls, what is to be held equal is our *basic liberties* – liberties that make democracy and an open society possible – such as speech, conscience, political participation, personal liberty, and so on (61). Rawls thinks that none of us would be willing to give these up under any normal circumstances, and nor would we be willing to have less of them than others.

But not everything is a basic liberty, and many things may be distributed unequally. For instance, jobs and positions of privilege are inevitably unequally distributed, and Rawls also thinks it would make sense to distribute goods like money unequally, provided the unequal distribution was good for everyone. The question, though, is what kinds and degrees of inequality are acceptable? Rawls’s second principle answers this by saying that inequalities of jobs and social status are acceptable if they are distributed

under conditions of *fair equality of opportunity*; and resulting inequalities of resources like income and wealth are acceptable if they are distributed so as to achieve the *greatest benefit to the least well-off*.

What do these conditions mean? Most fundamentally, opportunities are ‘formally’ equal when no one is legally excluded from the job market or directly discriminated against; for instance, no rule forbids women from working in important jobs (72). But this is insufficient if women tend to be hindered by nonlegal social practices such as gendered childcare and homemaking responsibilities. *Fair equality of opportunity* takes account of such inequalities that arise from social patterns. We can see that opportunity meets this standard if, for instance, the likelihood that a young child will grow up to be wealthy, or a CEO, or a parliamentarian, is no different whether that child is male or female, white or Black, from a poor family or a rich family, and so on. If fair equality of opportunity exists, you would expect a thorough re-sorting of people across income strata with every generation, with no patterns of differential outcomes between any two salient social groups.

And what of these income strata? The second part of the second principle – known as the “Difference Principle” – allows (or more accurately, requires) socioeconomic inequalities just insofar as these are “to the greatest benefit of the least advantaged” (302). This is a strange phrase; it sounds like the poorest are at the same time the richest, which makes no sense. But what Rawls means by it is that the point of inequality is to create, as the saying goes, a rising tide that lifts all boats (298); and in the Original Position we would choose whatever “tide” would lift all boats the highest. This is what he means by “greatest benefit of the least advantaged”: if we compare all possible societies, we’d create the one where the poorest are best-off compared to their counterparts in all other societies. However, it’s important to remember that this principle of inequality is the third-ranked idea, behind the equal basic liberties principle and fair equality of opportunity. So although we would choose unequal outcomes in exchange for increased wealth for everyone, we would not allow inequalities to grow to the point that our children would suffer unequal opportunity or that our political and moral freedoms were compromised.

In sum, Rawls’s final statement of the two principles is as follows:

First Principle

Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.

Second Principle

Social and economic inequalities are to be arranged so that they are both:

- (a) [Difference Principle] To the greatest benefit of the least advantaged, consistent with the just savings principle, and
- (b) Attached to offices and positions open to all under conditions of fair equality of opportunity (302).

Shorthand: when in the Original Position, design a society with the assumption that your worst enemy is going to place you in it (152).

Individual and Society

To be more precise, in the Original Position you are not exactly designing a society; that is to be done by constitution-writers, legislators, and the people (201). Instead, in the Original Position you are choosing principles that you will then use to evaluate whether the “basic structure” of your society is just (7). This basic structure – pervasive institutions like economic, political, and legal systems, as well as the standard social forms like family units, gender norms, and whether racial or ethnic classifications exist or have any impact on how our lives go – sets the framework for social cooperation. So when we choose principles of justice, we must bear in mind that the shape of the basic institutional structure has permanent and unchosen effects on our life prospects. Income differentials do not occur in a vacuum; they are products of those systems of distribution. For instance if, due to high stress, poor nutrition, inadequate schools, and unavailable childcare, poor parents are typically less able to prepare their children for career success, then the basic structure is enshrining intergenerational class hierarchies. Fair equality of opportunity would require us to dismantle such intergenerational hierarchies, at least provided we could do so without stunting the equal basic liberties.

In a just society, our market income and personal wealth – and Rawls does think that justice requires economic markets with personal property rights – significantly reflect past free choices that we and others have made. Many philosophers – including “natural rights” theorists such as Locke, “classical liberals” such as Hayek, and defenders of “meritocracy” – think that the relationship between free choice and market outcomes is what justifies our holdings: in their view, I have a ‘natural right’ to these things I own because I got them by using my own body in ways that neither coerced nor cheated others; or I ‘deserve’ what I have as a reward for how hard I worked and how talented I am.

Rawls disagrees. What you have to offer, and what it’s worth, depends on what others have and need, what people who came before you have invested in you and others, and the plans and expectations that people developed in light of a just system of interactions. Rawls sees what you deserve as what you may legitimately expect, *given* such institutions. He agrees that we all have important interests in controlling personal property and in being recognized for our meritorious efforts. But how we should accommodate these interests, and how weighty they are relative to other interests, are *outputs* of the theory of justice, not inputs. Market inequalities therefore have a function: to incentivize people to generate a surplus that can produce benefits and ease burdens for all. Relatively rich people can be said to ‘deserve’ their high incomes only because those incomes are generated by responding effectively to this incentive structure in a just society – that is, a society that makes the worst-off as well-off as they can be. And the scope of property rights is not ‘natural’ but instead depends on what justice says about which things can be owned, in what ways.

But if we are deeply shaped by the basic structure, and even our capacity to put forth meritorious effort is partly due to those social institutions, does this mean that each of us is merely an avatar of the society? Rawls may be said to be following Aristotle, John Stuart Mill, or arguably Kant in attempting to strike a balance between viewing the society as productive of individuals, including even their psychology, and viewing individuals as morally and rationally free and independent. If the natural rights and equal opportunity theorists went too far in imagining individuals as independent, certain communitarians go too far in the opposite direction. Rawls views each of us as independent enough to form and revise our life plan and values, and thinks it is important that each of us be held responsible for

these life plans and values. Moreover, his Original Position is designed so that no one's basic interests are sacrificed for the benefit of others. But he also views us as shaped by, and typically finding our happiness in, collective endeavors and social systems. Moreover, how we develop our individual life plans is answerable to the principles of justice; notwithstanding my liberty of conscience, I may not create a religion based on enslavement or which forbids anyone from leaving the religious community; notwithstanding my freedom to form and revise my conception of the good, I can be blamed and punished if that conception of the good involves violating others' rights. For Rawls, the good society is not a "community," because it is not unified by a shared conception of the good. But the good society *is* a "social union," because it is a freely affirmed cooperative arrangement unified by justice and enabling us to value and benefit from one another's excellences (523).

Applied and 'Non-Ideal' Political Philosophy

A Theory of Justice has had major impacts in ethics and other areas of philosophy, and inspired a renaissance of Kant scholarship and broadly Kantian methods through the entire discipline of Philosophy. But of course its greatest impact was in political philosophy, a field that was reshaped and reinvigorated – arguably worldwide, but certainly in the Anglo-American tradition – from the moment *TJ* appeared. Many scholars began working within the new set of frameworks and questions that *TJ* foregrounded. It was the lodestar of the field.

So it is puzzling, to say the least, that *TJ* virtually ignores all of the most urgent problems of social justice in the world around it. How could the central work of normative political philosophy, the greatest achievement of the theory of social justice to emerge from academic Philosophy in the West, say virtually nothing about racial oppression, migration, international justice, gender, the environment, or the legacies of colonialism and imperialism? Notwithstanding Rawls's inclusion of brief chapters on civil disobedience, just war, and intergenerational justice, it is hard to believe this work is a product of the 1960s.

Some of *TJ*'s limitations may be endemic to "analytic" political philosophy, specifically in the liberal or social-contract tradition – for instance, the emphasis on individuals, law, and the methods and assumptions of neoclassical economics – and the consequent difficulty of addressing extralegal social power, racial formation, gender norms, the Earth's finitude, and so on. Other limits may be due to Rawls's decision to focus on "ideal theory" – rather than the more urgent task of "non-ideal" theory – on the assumption that you need to know what target to aim at in order to know which social changes would move you towards rather than away from that target. And still other limitations may be due to the demographics and social structure of the profession of philosophy at the time, namely, that it was largely an in-club of financially comfortable white men at a handful of elite institutions who mostly read one another's work.

A wide variety of these issues have been addressed in the philosophical literature of recent decades. A concluding task for this review is to discuss how Rawls's brief remarks on issues of "applied" political philosophy, plus his philosophical methodology, did or could address these ignored questions. It is not claimed that *TJ*'s methods would be successful in addressing them; debate in every area continues.

An early challenge to *TJ* had to do with international justice. Rawls imagined that the society in question was "closed," meaning that people enter only by birth and exit only by death, and the principles of justice apply only within such societies and not beyond them. This seems to imply that Rawls has no objection to staggering inequality across borders, or even colonial relations between states, just so long

as there is a liberal democracy within each state. (Rawls's later foray into international justice – *The Law of Peoples* – did not do much to allay these concerns, but opened new lines of criticism.) Moreover, scholars of migration have wondered how he could be so ignorant of the perennial fact that people cross borders, particularly given that the US is widely seen as a “nation of immigrants.”

Rawls would deny that his “closed” society assumption has to do with migration at all. Rather, it reflects Rawls's view of how the principles are to be chosen and justified. Those who don the Veil of Ignorance must choose institutions they would be willing to live with, *come what may*. Suppose you were in the Original Position and you thought, “If I come out on top, I'll stay, but if not, I'll leave.” You might then take a gamble, creating a society with higher peaks and lower valleys. But worse, you would then be choosing not for yourself but for others: for those left behind or unable to leave. Thus although it makes it difficult to discuss justice in the context of migration and bordering practices, Rawls's “closed society” assumption was an attempt to ensure that principles of justice are self-imposed and hence freely affirmed by each of us, and that we choose only something that we can live with. As for international or global justice, many commentators have argued that – notwithstanding what Rawls himself says in *TJ* or later in *Law of Peoples* – his methodology does justify a considerably more robust system of global justice (e.g. Beitz). However, Rawls remained concerned to enable moral, political, and cultural diversity, and worried that a uniform global basic structure would require a degree of socioeconomic and political integration that would violate the equal basic liberties and fair equality of opportunity principles. He thus endorsed a “global Society of Peoples” instead of affirming a uniform political and economic system subject to the principles of justice.

Similarly, early critics such as Jane English raised the problem of intergenerational justice and the global environment. To address this question Rawls prescribes that the Veil of Ignorance hide from us which generation we are born into and how wealthy our society is. He thinks we would choose a “Just Savings” principle requiring that each generation play its part in the multi-generational enterprise of achieving and maintaining just institutions. In later work, Rawls revised this in line with English's suggestion, so that we would know our generation but choose a principle that we would want all previous generations to have followed. It is unclear whether either formulation of the just savings principle could achieve the unanimity across generations that Rawls's method seems to require in order to produce determinate results. Nor does either method enfranchise animals or the natural environment, thus ensuring that Rawls's system remains anthropocentric. In addition to environmental challenges that arise due to the Earth's *finitude*, other environmental challenges arise due to the Earth's *inevitable unpredictability*: earthquakes, disasters, pandemics, and so on. In the Original Position we should know that such disasters are an inevitable part of life on Earth, and we would presumably demand that just institutions be structured for ethical disaster preparedness, mitigation, and burden-sharing. It is not clear, however, where if at all such a principle fits within Rawls's two principles.

Gender, Race, and Colonialism

One jarring feature of Rawls's discussion of intergenerational justice is that he frames it in terms of what fathers owe to sons and what people in the Original Position would choose if they imagined themselves as fathers. Further, in addressing the problem of just savings he recognizes that justice is a multigenerational enterprise, but he never considers what those who enjoy more-just institutions owe to the descendants of those who, in the past, were treated unjustly. These two gaps open out to major questions about gender justice and racial justice both within and across generations.

Numerous scholars, most famously Susan Moller Okin, have argued that, try as we might, it will be impossible to repair Rawls's methodology to address justice between genders or within the family. One way of seeing the problem is by noting that households have different numbers of people in them; some adults live alone, others live in households of one or more other wage-earners and multiple children and retirees, and so on. Is the degree of social inequality, then, to be measured across households, or across individuals? Suppose distributive justice allows a maximum wealth difference of 5:1 between the richest and the poorest individuals. But then, two rich people marry and found a family. Suddenly the rich household has 10 times the wealth of the poorest individual. So the whole system is put off. But if the permissible wealth differentials are between *households*, now the principles have nothing to say about justice between spouses. Similar problems arise when we think of justice between *overlapping* generations. So far, at least in modern societies, humans do not seem to have invented a social unit, other than the family, in which children can be reared with the love and attention they need. Yet if in a just society the social mores and habits inculcated in better-off households enhance their children's competitive chances in the economy going forward, the sorting required for fair equality of opportunity will be hindered, and class status will be heritable (74). The family is thus both required for justice, and corrosive of it.

Now imagine a society that meets the standards of Rawlsian justice in the abstract, but women are disproportionately at the low end while men are disproportionately at the high end of the distributive scheme; or racial minorities disproportionately low and white people disproportionately high. Has anything gone wrong, by the lights of Rawls's two principles? Further, suppose that one reason for these unequal outcomes is that young women or minorities in the education system tend to believe that they are not "cut out for" the most lucrative or powerful sectors, and hence form preferences to pursue lower-status and lower-paying positions, leading to a sorting that seems both unjust and yet perfectly voluntary. Again: are Rawls's two principles any help here?

Rawls suggests that, once out from under the Veil of Ignorance, we should be able to take any salient social identity and check it for these kinds of systematic biases, and continually refine the "precepts of justice" within the society in order to achieve and maintain background justice (307). But it will be a challenge to fix these problems; liberal principles have a hard time addressing 'unofficial' social mores, behavior patterns, and sedimented gender roles – all the more so if we have not worked out the intrafamilial justice issues mentioned above. Even deeper problems arise in the context of historic injustices such as colonialism, slavery, segregation, and the legacy of periodic flareups of intercommunal terrorism used to keep these systems in place. If property rights or social capital are in any way heritable, shouldn't descendants of victims now be compensated by descendants of perpetrators? Could a principle of intergenerational justice include a requirement to make whole, in some way, the past victims of the unjust practices of the present state? And even more fundamentally: if we live in a settler-colonial state, who should be choosing principles of justice for us – the wrongfully displaced indigenous peoples and their descendants, the colonizers and their descendants, or all together? Whichever answer we give here, we seem to be choosing a side in a debate that is itself about justice, and so the theory of justice presupposes prior principles of just territorial control or what it is for us all to be one people.

Other Applied Issues

Many people, philosophers and others, find Rawls's basic methodology compelling, and agree that his principles articulate, if not fully sufficient conditions of justice, at least good shorthand starting points. Consequently, the Original Position model with a Veil of Ignorance has become a very familiar and

popular method of posing questions and framing theories in virtually every social context where benefits and burdens are distributed, including health and health care (Daniels), the workplace (Arnold), environmental justice (Bell), and racial justice (Shelby). Though the *direct* impact of *Theory of Justice* may have waned, its indirect impacts continue to expand.

Questions for Self-Review

1. Rawls calls his view “justice as fairness” because it is the product of free choice in a fair choice situation. What makes the choice situation “fair”?
2. Rawls prioritizes the “equal basic liberties,” emphasizing the plural (a whole system of liberties rather than any single liberty). If you think your basic *rights* are important, is the priority of liberties sufficient to protect basic rights? What would be different if Rawls had made basic *rights* primary instead?
3. Rawls remarked that his principles of justice can be said to capture the classic liberal ideals of *liberty, equality, and fraternity*, with the Difference Principle specifically connecting to the ideal of fraternity. How do the principles capture these ideals?
4. Cardiologists – heart doctors – have the highest median pay of any career in the United States. Do they deserve their high pay? How would Rawls assess this question?
5. Choose some area of contemporary struggle over social justice. How might Rawls’s methodology be applied to this question?

Works Cited & Supplemental Reading

- Arnold, Samuel. “The Difference Principle at Work.” *Journal of Political Philosophy* 20 #1 (2012), 94-118.
- Beitz, Charles. *Political Theory and International Relations*, 2nd ed. Princeton: Princeton University Press, 1999.
- Bell, Derek. “Environmental Justice and Rawls’s Difference Principle.” *Environmental Ethics* 26 #3 (2004), 287-306.
- Daniels, Norman. *Just Health*. New York: Cambridge University Press, 2007.
- English, Jane. “Justice Between Generations.” *Philosophical Studies* 31 #1 (1977), 91-104.
- Mills, Charles W. *Black Rights/White Wrongs*. New York: Oxford University Press, 2017.
- Okin, Susan Moller. *Justice, Gender, and the Family*. New York: Basic Books, 1991.
- Rawls, John. *Justice as Fairness: A Restatement*, ed. Erin Kelly. Cambridge, Mass: Belknap Press/Harvard University Press, 2001.
- Rawls, John. *The Law of Peoples*. Cambridge, Mass: Harvard University Press, 1999.
- Shelby, Tommie. *Dark Ghettos: Injustice, Dissent, and Reform*. Cambridge, Mass: Belknap Press/Harvard University Press, 2018.