

[INTERNET ENCYCLOPEDIA OF PHILOSOPHY](#)

[Social Contract Theory](#)

Social contract theory, nearly as old as philosophy itself, is the view that persons' moral and/or political obligations are dependent upon a contract or agreement among them to form the society in which they live. Socrates uses something quite like a social contract argument to explain to Crito why he must remain in prison and accept the death penalty. However, social contract theory is rightly associated with modern moral and political theory and is given its first full exposition and defense by Thomas Hobbes. After Hobbes, John Locke and Jean-Jacques Rousseau are the best known proponents of this enormously influential theory, which has been one of the most dominant theories within moral and political theory throughout the history of the modern West. In the twentieth century, moral and political theory regained philosophical momentum as a result of John Rawls' Kantian version of social contract theory, and was followed by new analyses of the subject by David Gauthier and others. More recently, philosophers from different perspectives have offered new criticisms of social contract theory. In particular, feminists and race-conscious philosophers have argued that social contract theory is at least an incomplete picture of our moral and political lives, and may in fact camouflage some of the ways in which the contract is itself parasitical upon the subjugations of classes of persons.

Table of Contents

1. [Socrates' Argument](#)
2. [Modern Social Contract Theory](#)
 1. [Thomas Hobbes](#)
 2. [John Locke](#)
 3. [Jean-Jacques Rousseau](#)
3. [More Recent Social Contract Theories](#)
 1. [John Rawls' *A Theory of Justice*](#)
 2. [David Gauthier](#)
4. [Contemporary Critiques of Social Contract Theory](#)
 1. [Feminist Arguments](#)
 1. [The Sexual Contract](#)
 2. [The Nature of the Liberal Individual](#)

3. [Arguing from Care](#)
2. [Race-Conscious Argument](#)
5. [Conclusion](#)
6. [References and Further Reading](#)

1. Socrates' Argument

In the early Platonic dialogue, *Crito*, [Socrates](#) makes a compelling argument as to why he must stay in prison and accept the death penalty, rather than escape and go into exile in another Greek city. He personifies the Laws of Athens, and, speaking in their voice, explains that he has acquired an overwhelming obligation to obey the Laws because they have made his entire way of life, and even the fact of his very existence, possible. They made it possible for his mother and father to marry, and therefore to have legitimate children, including himself. Having been born, the city of Athens, through its laws, then required that his father care for and educate him. Socrates' life and the way in which that life has flourished in Athens are each dependent upon the Laws. Importantly, however, this relationship between citizens and the Laws of the city are not coerced. Citizens, once they have grown up, and have seen how the city conducts itself, can choose whether to leave, taking their property with them, or stay. Staying implies an agreement to abide by the Laws and accept the punishments that they mete out. And, having made an agreement that is itself just, Socrates asserts that he must keep to this agreement that he has made and obey the Laws, in this case, by staying and accepting the death penalty. Importantly, the contract described by Socrates is an implicit one: it is implied by his choice to stay in Athens, even though he is free to leave.

In Plato's most well-known dialogue, *Republic*, social contract theory is represented again, although this time less favorably. In Book II, Glaucon offers a candidate for an answer to the question "what is justice?" by representing a social contract explanation for the nature of justice. What men would most want is to be able to commit injustices against others without the fear of reprisal, and what they most want to avoid is being treated unjustly by others without being able to do injustice in return. Justice then, he says, is the conventional result of the laws and covenants that men make in order to avoid these extremes. Being unable to commit injustice with impunity (as those who wear the ring of Gyges would), and fearing becoming victims themselves, men decide that it is in their interests to submit themselves to the convention of justice. Socrates rejects this view, and most of the rest of the dialogue centers on showing that justice is worth having for its own sake, and that

the just man is the happy man. So, from Socrates' point of view, justice has a value that greatly exceeds the prudential value that Glaucon assigns to it.

These views, in the *Crito* and the *Republic*, might seem at first glance inconsistent: in the former dialogue Socrates uses a social contract type of argument to show why it is just for him to remain in prison, whereas in the latter he rejects social contract as the source of justice. These two views are, however, reconcilable. From Socrates' point of view, a just man is one who will, among other things, recognize his obligation to the state by obeying its laws. The state is the morally and politically most fundamental entity, and as such deserves our highest allegiance and deepest respect. Just men know this and act accordingly. Justice, however, is more than simply obeying laws in exchange for others obeying them as well. Justice is the state of a well-regulated soul, and so the just man will also necessarily be the happy man. So, justice is more than the simple reciprocal obedience to law, as Glaucon suggests, but it does nonetheless include obedience to the state and the laws that sustain it. So in the end, although Plato is perhaps the first philosopher to offer a representation of the argument at the heart of social contract theory, Socrates ultimately rejects the idea that social contract is the original source of justice.

2. Modern Social Contract Theory

a. Thomas Hobbes

[Thomas Hobbes](#), 1588-1679, lived during the most crucial period of early modern England's history: the English Civil War, waged from 1642-1648. To describe this conflict in the most general of terms, it was a clash between the King and his supporters, the Monarchists, who preferred the traditional authority of a monarch, and the Parliamentarians, most notably led by Oliver Cromwell, who demanded more power for the quasi-democratic institution of Parliament. Hobbes represents a compromise between these two factions. On the one hand he rejects the theory of the Divine Right of Kings, which is most eloquently expressed by Robert Filmer in his *Patriarcha or the Natural Power of Kings*, (although it would be left to John Locke to refute Filmer directly). Filmer's view held that a king's authority was invested in him (or, presumably, her) by God, that such authority was absolute, and therefore that the basis of political obligation lay in our obligation to obey God absolutely. According to this view, then, political obligation is subsumed under religious obligation. On the other hand, Hobbes also rejects the early democratic view, taken up by the Parliamentarians, that power ought to be shared between

Parliament and the King. In rejecting both these views, Hobbes occupies the ground of one who is both radical and conservative. He argues, radically for his times, that political authority and obligation are based on the individual self-interests of members of society who are understood to be equal to one another, with no single individual invested with any essential authority to rule over the rest, while at the same time maintaining the conservative position that the monarch, which he called the Sovereign, must be ceded absolute authority if society is to survive.

Hobbes' political theory is best understood if taken in two parts: his theory of human motivation, [Psychological Egoism](#), and his theory of the social contract, founded on the hypothetical State of Nature. Hobbes has, first and foremost, a particular theory of human nature, which gives rise to a particular view of morality and politics, as developed in his philosophical masterpiece, *Leviathan*, published in 1651. The Scientific Revolution, with its important new discoveries that the universe could be both described and predicted in accordance with universal laws of nature, greatly influenced Hobbes. He sought to provide a theory of human nature that would parallel the discoveries being made in the sciences of the inanimate universe. His psychological theory is therefore informed by mechanism, the general view that everything in the universe is produced by nothing other than matter in motion. According to Hobbes, this extends to human behavior. Human macro-behavior can be aptly described as the effect of certain kinds of micro-behavior, even though some of this latter behavior is invisible to us. So, such behaviors as walking, talking, and the like are themselves produced by other actions inside of us. And these other actions are themselves caused by the interaction of our bodies with other bodies, human or otherwise, which create in us certain chains of causes and effects, and which eventually give rise to the human behavior that we can plainly observe. We, including all of our actions and choices, are then, according to this view, as explainable in terms of universal laws of nature as are the motions of heavenly bodies. The gradual disintegration of memory, for example, can be explained by inertia. As we are presented with ever more sensory information, the residue of earlier impressions 'slows down' over time. From Hobbes' point of view, we are essentially very complicated organic machines, responding to the stimuli of the world mechanistically and in accordance with universal laws of human nature.

In Hobbes' view, this mechanistic quality of human psychology implies the subjective nature of normative claims. 'Love' and 'hate', for instance, are just words we use to describe the things we are drawn to and repelled by, respectively. So,

too, the terms 'good' and 'bad' have no meaning other than to describe our appetites and aversions. Moral terms do not, therefore, describe some objective state of affairs, but are rather reflections of individual tastes and preferences.

In addition to Subjectivism, Hobbes also infers from his mechanistic theory of human nature that humans are necessarily and exclusively self-interested. All men pursue only what they perceive to be in their own individually considered best interests – they respond mechanistically by being drawn to that which they desire and repelled by that to which they are averse. This is a universal claim: it is meant to cover all human actions under all circumstances – in society or out of it, with regard to strangers and friends alike, with regard to small ends and the most generalized of human desires, such as the desire for power and status. Everything we do is motivated solely by the desire to better our own situations, and satisfy as many of our own, individually considered desires as possible. We are infinitely appetitive and only genuinely concerned with our own selves. According to Hobbes, even the reason that adults care for small children can be explicated in terms of the adults' own self-interest (he claims that in saving an infant by caring for it, we become the recipient of a strong sense of obligation in one who has been helped to survive rather than allowed to die).

In addition to being exclusively self-interested, Hobbes also argues that human beings are reasonable. They have in them the rational capacity to pursue their desires as efficiently and maximally as possible. Their reason does not, given the subjective nature of value, evaluate their given ends, rather it merely acts as "Scouts, and Spies, to range abroad, and find the way to the things Desired" (139). Rationality is purely instrumental. It can add and subtract, and compare sums one to another, and thereby endows us with the capacity to formulate the best means to whatever ends we might happen to have.

From these premises of human nature, Hobbes goes on to construct a provocative and compelling argument for why we ought to be willing to submit ourselves to political authority. He does this by imagining persons in a situation prior to the establishment of society, the State of Nature.

According to Hobbes, the justification for political obligation is this: given that men are naturally self-interested, yet they are rational, they will choose to submit to the authority of a Sovereign in order to be able to live in a civil society, which is conducive to their own interests. Hobbes argues for this by imagining men in their

natural state, or in other words, the State of Nature. In the State of Nature, which is purely hypothetical according to Hobbes, men are naturally and exclusively self-interested, they are more or less equal to one another, (even the strongest man can be killed in his sleep), there are limited resources, and yet there is no power able to force men to cooperate. Given these conditions in the State of Nature, Hobbes concludes that the State of Nature would be unbearably brutal. In the State of Nature, every person is always in fear of losing his life to another. They have no capacity to ensure the long-term satisfaction of their needs or desires. No long-term or complex cooperation is possible because the State of Nature can be aptly described as a state of utter distrust. Given Hobbes' reasonable assumption that most people want first and foremost to avoid their own deaths, he concludes that the State of Nature is the worst possible situation in which men can find themselves. It is the state of perpetual and unavoidable war.

The situation is not, however, hopeless. Because men are reasonable, they can see their way out of such a state by recognizing the laws of nature, which show them the means by which to escape the State of Nature and create a civil society. The first and most important law of nature commands that each man be willing to pursue peace when others are willing to do the same, all the while retaining the right to continue to pursue war when others do not pursue peace. Being reasonable, and recognizing the rationality of this basic precept of reason, men can be expected to construct a Social Contract that will afford them a life other than that available to them in the State of Nature. This contract is constituted by two distinguishable contracts. First, they must agree to establish society by collectively and reciprocally renouncing the rights they had against one another in the State of Nature. Second, they must imbue some one person or assembly of persons with the authority and power to enforce the initial contract. In other words, to ensure their escape from the State of Nature, they must both agree to live together under common laws, and create an *enforcement mechanism* for the social contract and the laws that constitute it. Since the sovereign is invested with the authority and power to mete out punishments for breaches of the contract which are worse than not being able to act as one pleases, men have good, albeit self-interested, reason to adjust themselves to the artifice of morality in general, and justice in particular. Society becomes possible because, whereas in the State of Nature there was no power able to "overawe them all", now there is an artificially and conventionally superior and more powerful person who can force men to cooperate. While living under the authority of a Sovereign can be harsh (Hobbes argues that because men's passions can be expected to overwhelm their reason, the Sovereign must have

absolute authority in order for the contract to be successful) it is at least better than living in the State of Nature. And, no matter how much we may object to how poorly a Sovereign manages the affairs of the state and regulates our own lives, we are never justified in resisting his power because it is the only thing which stands between us and what we most want to avoid, the State of Nature.

According to this argument, morality, politics, society, and everything that comes along with it, all of which Hobbes calls 'commodious living' are purely conventional. Prior to the establishment of the basic social contract, according to which men agree to live together and the contract to embody a Sovereign with absolute authority, nothing is immoral or unjust – anything goes. After these contracts are established, however, then society becomes possible, and people can be expected to keep their promises, cooperate with one another, and so on. The Social Contract is the most fundamental source of all that is good and that which we depend upon to live well. Our choice is either to abide by the terms of the contract, or return to the State of Nature, which Hobbes argues no reasonable person could possibly prefer.

Given his rather severe view of human nature, Hobbes nonetheless manages to create an argument that makes civil society, along with all its advantages, possible. Within the context of the political events of his England, he also managed to argue for a continuation of the traditional form of authority that his society had long since enjoyed, while nonetheless placing it on what he saw as a far more acceptable foundation.

b. John Locke

For Hobbes, the necessity of an absolute authority, in the form of a Sovereign, followed from the utter brutality of the State of Nature. The State of Nature was completely intolerable, and so rational men would be willing to submit themselves even to absolute authority in order to escape it. For [John Locke](#), 1632-1704, the State of Nature is a very different type of place, and so his argument concerning the social contract and the nature of men's relationship to authority are consequently quite different. While Locke uses Hobbes' methodological device of the State of Nature, as do virtually all social contract theorists, he uses it to a quite different end. Locke's arguments for the social contract, and for the right of citizens to revolt against their king were enormously influential on the democratic

revolutions that followed, especially on Thomas Jefferson, and the founders of the United States.

Locke's most important and influential political writings are contained in his *Two Treatises on Government*. The first treatise is concerned almost exclusively with refuting the argument of Robert Filmer's *Patriarcha*, that political authority was derived from religious authority, also known by the description of the Divine Right of Kings, which was a very dominant theory in seventeenth-century England. The second treatise contains Locke's own constructive view of the aims and justification for civil government, and is titled "An Essay Concerning the True Original Extent and End of Civil Government".

According to Locke, the State of Nature, the natural condition of mankind, is a state of perfect and complete liberty to conduct one's life as one best sees fit, free from the interference of others. This does not mean, however, that it is a state of license: one is not free to do anything at all one pleases, or even anything that one judges to be in one's interest. The State of Nature, although a state wherein there is no civil authority or government to punish people for transgressions against laws, is not a state without morality. The State of Nature is pre-political, but it is not pre-moral. Persons are assumed to be equal to one another in such a state, and therefore equally capable of discovering and being bound by the Law of Nature. The Law of Nature, which is on Locke's view the basis of all morality, and given to us by God, commands that we not harm others with regards to their "life, health, liberty, or possessions" (par. 6). Because we all belong equally to God, and because we cannot take away that which is rightfully His, we are prohibited from harming one another. So, the State of Nature is a state of liberty where persons are free to pursue their own interests and plans, free from interference, and, because of the Law of Nature and the restrictions that it imposes upon persons, it is relatively peaceful.

The State of Nature therefore, is not the same as the state of war, as it is according to Hobbes. It can, however devolve into a state of war, in particular, a state of war over property disputes. Whereas the State of Nature is the state of liberty where persons recognize the Law of Nature and therefore do not harm one another, the state of war begins between two or more men once one man declares war on another, by stealing from him, or by trying to make him his slave. Since in the State of Nature there is no civil power to whom men can appeal, and since the Law of Nature allows them to defend their own lives, they may then kill those who would

bring force against them. Since the State of Nature lacks civil authority, once war begins it is likely to continue. And this is one of the strongest reasons that men have to abandon the State of Nature by contracting together to form civil government.

Property plays an essential role in Locke's argument for civil government and the contract that establishes it. According to Locke, private property is created when a person mixes his labor with the raw materials of nature. So, for example, when one tills a piece of land in nature, and makes it into a piece of farmland, which produces food, then one has a claim to own that piece of land and the food produced upon it. (This led Locke to conclude that America didn't really belong to the natives who lived there, because they were, on his view, failing to utilize the basic material of nature. In other words, they didn't farm it, so they had no legitimate claim to it, and others could therefore justifiably appropriate it.) Given the implications of the Law of Nature, there are limits as to how much property one can own: one is not allowed to take more from nature than one can use, thereby leaving others without enough for themselves. Because nature is given to all of mankind by God for its common subsistence, one cannot take more than his own fair share. Property is the linchpin of Locke's argument for the social contract and civil government because it is the protection of their property, including their property in their own bodies, that men seek when they decide to abandon the State of Nature.

According to Locke, the State of Nature is not a condition of individuals, as it is for Hobbes. Rather, it is populated by mothers and fathers with their children, or families – what he calls “conjugal society” (par. 78). These societies are based on the voluntary agreements to care for children together, and they are moral but not political. Political society comes into being when individual men, representing their families, come together in the State of Nature and agree to each give up the executive power to punish those who transgress the Law of Nature, and hand over that power to the public power of a government. Having done this, they then become subject to the will of the majority. In other words, by making a compact to leave the State of Nature and form society, they make “one body politic under one government” (par. 97) and submit themselves to the will of that body. One joins such a body, either from its beginnings, or after it has already been established by others, only by explicit consent. Having created a political society and government through their consent, men then gain three things which they lacked in the State of Nature: laws, judges to adjudicate laws, and the executive power necessary to enforce these laws. Each man therefore gives over the power to protect himself

and punish transgressors of the Law of Nature to the government that he has created through the compact.

Given that the end of “men’s uniting into common-wealths”(par. 124) is the preservation of their wealth, and preserving their lives, liberty, and well-being in general, Locke can easily imagine the conditions under which the compact with government is destroyed, and men are justified in resisting the authority of a civil government, such as a King. When the executive power of a government devolves into tyranny, such as by dissolving the legislature and therefore denying the people the ability to make laws for their own preservation, then the resulting tyrant puts himself into a State of Nature, and specifically into a state of war with the people, and they then have the same right to self-defense as they had before making a compact to establish society in the first place. In other words, the justification of the authority of the executive component of government is the protection of the people’s property and well-being, so when such protection is no longer present, or when the king becomes a tyrant and acts against the interests of the people, they have a right, if not an outright obligation, to resist his authority. The social compact can be dissolved and the process to create political society begun anew.

Because Locke did not envision the State of Nature as grimly as did Hobbes, he can imagine conditions under which one would be better off rejecting a particular civil government and returning to the State of Nature, with the aim of constructing a better civil government in its place. It is therefore both the view of human nature, and the nature of morality itself, which account for the differences between Hobbes’ and Locke’s views of the social contract.

c. Jean-Jacques Rousseau

[Jean-Jacques Rousseau](#), 1712-1778, lived and wrote during what was arguably the headiest period in the intellectual history of modern France—the Enlightenment. He was one of the bright lights of that intellectual movement, contributing articles to the *Encyclopedie* of Diderot, and participating in the salons in Paris, where the great intellectual questions of his day were pursued.

Rousseau has two distinct social contract theories. The first is found in his essay, *Discourse on the Origin and Foundations of Inequality Among Men*, commonly referred to as the Second Discourse, and is an account of the moral and political evolution of human beings over time, from a State of Nature to modern society. As

such it contains his *naturalized* account of the social contract, which he sees as very problematic. The second is his *normative*, or idealized theory of the social contract, and is meant to provide the means by which to alleviate the problems that modern society has created for us, as laid out in the *Social Contract*.

Rousseau wrote his *Second Discourse* in response to an essay contest sponsored by the Academy of Dijon. (Rousseau had previously won the same essay contest with an earlier essay, commonly referred to as the *First Discourse*.) In it he describes the historical process by which man began in a State of Nature and over time 'progressed' into civil society. According to Rousseau, the State of Nature was a peaceful and quixotic time. People lived solitary, uncomplicated lives. Their few needs were easily satisfied by nature. Because of the abundance of nature and the small size of the population, competition was non-existent, and persons rarely even saw one another, much less had reason for conflict or fear. Moreover, these simple, morally pure persons were naturally endowed with the capacity for pity, and therefore were not inclined to bring harm to one another.

As time passed, however, humanity faced certain changes. As the overall population increased, the means by which people could satisfy their needs had to change. People slowly began to live together in small families, and then in small communities. Divisions of labor were introduced, both within and between families, and discoveries and inventions made life easier, giving rise to leisure time. Such leisure time inevitably led people to make comparisons between themselves and others, resulting in public values, leading to shame and envy, pride and contempt. Most importantly however, according to Rousseau, was the invention of private property, which constituted the pivotal moment in humanity's evolution out of a simple, pure state into one characterized by greed, competition, vanity, inequality, and vice. For Rousseau the invention of property constitutes humanity's 'fall from grace' out of the State of Nature.

Having introduced private property, initial conditions of inequality became more pronounced. Some have property and others are forced to work for them, and the development of social classes begins. Eventually, those who have property notice that it would be in their interests to create a government that would protect private property from those who do not have it but can see that they might be able to acquire it by force. So, government gets established, through a contract, which purports to guarantee equality and protection for all, even though its true purpose is to fossilize the very inequalities that private property has produced. In other

words, the contract, which claims to be in the interests of everyone equally, is really in the interests of the few who have become stronger and richer as a result of the developments of private property. This is the naturalized social contract, which Rousseau views as responsible for the conflict and competition from which modern society suffers.

The normative social contract, argued for by Rousseau in *The Social Contract* (1762), is meant to respond to this sorry state of affairs and to remedy the social and moral ills that have been produced by the development of society. The distinction between history and justification, between the factual situation of mankind and how it ought to live together, is of the utmost importance to Rousseau. While we ought not to ignore history, nor ignore the causes of the problems we face, we must resolve those problems through our capacity to choose how we ought to live. Might never makes right, despite how often it pretends that it can.

The Social Contract begins with the most oft-quoted line from Rousseau: “Man was born free, and he is everywhere in chains” (49). This claim is the conceptual bridge between the descriptive work of the Second Discourse, and the prescriptive work that is to come. Humans are essentially free, and were free in the State of Nature, but the ‘progress’ of civilization has substituted subservience to others for that freedom, through dependence, economic and social inequalities, and the extent to which we judge ourselves through comparisons with others. Since a return to the State of Nature is neither feasible nor desirable, the purpose of politics is to restore freedom to us, thereby reconciling who we truly and essentially are with how we live together. So, this is the fundamental philosophical problem that *The Social Contract* seeks to address: how can we be free and live together? Or, put another way, how can we live together without succumbing to the force and coercion of others? We can do so, Rousseau maintains, by submitting our individual, particular wills to the collective or general will, created through agreement with other free and equal persons. Like Hobbes and Locke before him, and in contrast to the ancient philosophers, all men are made by nature to be equals, therefore no one has a natural right to govern others, and therefore the only justified authority is the authority that is generated out of agreements or covenants.

The most basic covenant, the social pact, is the agreement to come together and form a people, a collectivity, which by definition is more than and different from a mere aggregation of individual interests and wills. This act, where individual

persons become a people is “the real foundation of society” (59). Through the collective renunciation of the individual rights and freedom that one has in the State of Nature, and the transfer of these rights to the collective body, a new ‘person’, as it were, is formed. The sovereign is thus formed when free and equal persons come together and agree to create themselves anew as a single body, directed to the good of all considered together. So, just as individual wills are directed towards individual interests, the general will, once formed, is directed towards the common good, understood and agreed to collectively. Included in this version of the social contract is the idea of reciprocated duties: the sovereign is committed to the good of the individuals who constitute it, and each individual is likewise committed to the good of the whole. Given this, individuals cannot be given liberty to decide whether it is in their own interests to fulfill their duties to the Sovereign, while at the same time being allowed to reap the benefits of citizenship. They must be made to conform themselves to the general will, they must be “forced to be free” (64).

For Rousseau, this implies an extremely strong and direct form of democracy. One cannot transfer one’s will to another, to do with as he or she sees fit, as one does in representative democracies. Rather, the general will depends on the coming together periodically of the entire democratic body, each and every citizen, to decide collectively, and with at least near unanimity, how to live together, i.e., what laws to enact. As it is constituted only by individual wills, these private, individual wills must assemble themselves regularly if the general will is to continue. One implication of this is that the strong form of democracy which is consistent with the general will is also only possible in relatively small states. The people must be able to identify with one another, and at least know who each other are. They cannot live in a large area, too spread out to come together regularly, and they cannot live in such different geographic circumstances as to be unable to be united under common laws. (Could the present-day U.S. satisfy Rousseau’s conception of democracy? It could not.) Although the conditions for true democracy are stringent, they are also the only means by which we can, according to Rousseau, save ourselves, and regain the freedom to which we are naturally entitled.

Rousseau’s social contract theories together form a single, consistent view of our moral and political situation. We are endowed with freedom and equality by nature, but our nature has been corrupted by our contingent social history. We can overcome this corruption, however, by invoking our free will to reconstitute

ourselves politically, along strongly democratic principles, which is good for us, both individually and collectively.

3. More Recent Social Contract Theories

a. John Rawls' A Theory of Justice

In 1972, the publication of [John Rawls'](#) extremely influential *A Theory of Justice* brought moral and political philosophy back from what had been a long hiatus of philosophical consideration. Rawls' theory relies on a Kantian understanding of persons and their capacities. For Rawls, as for Kant, persons have the capacity to reason from a universal point of view, which in turn means that they have the particular moral capacity of judging principles from an impartial standpoint. In *A Theory of Justice*, Rawls argues that the moral and political point of view is discovered via impartiality. (It is important to note that this view, delineated in *A Theory of Justice*, has undergone substantial revisions by Rawls, and that he described his later view as "political liberalism".) He invokes this point of view (the general view that Thomas Nagel describes as "the view from nowhere") by imagining persons in a hypothetical situation, the Original Position, which is characterized by the epistemological limitation of the Veil of Ignorance. Rawls' original position is his highly abstracted version of the State of Nature. It is the position from which we can discover the nature of justice and what it requires of us as individual persons and of the social institutions through which we will live together cooperatively. In the original position, behind the veil of ignorance, one is denied any particular knowledge of one's circumstances, such as one's gender, race, particular talents or disabilities, one's age, social status, one's particular conception of what makes for a good life, or the particular state of the society in which one lives. Persons are also assumed to be rational and disinterested in one another's well-being. These are the conditions under which, Rawls argues, one can choose principles for a just society which are themselves chosen from initial conditions that are inherently fair. Because no one has any of the particular knowledge he or she could use to develop principles that favor his or her own particular circumstances, in other words the knowledge that makes for and sustains prejudices, the principles chosen from such a perspective are necessarily fair. For example, if one does not know whether one is female or male in the society for which one must choose basic principles of justice, it makes no sense, from the point of view of self-interested rationality, to endorse a principle that favors one sex at the expense of another, since, once the veil of ignorance is lifted, one might find

oneself on the losing end of such a principle. Hence Rawls describes his theory as “justice as fairness.” Because the conditions under which the principles of justice are discovered are basically fair, justice proceeds out of fairness.

In such a position, behind such a veil, everyone is in the same situation, and everyone is presumed to be equally rational. Since everyone adopts the same method for choosing the basic principles for society, everyone will occupy the same standpoint: that of the disembodied, rational, universal human. Therefore all who consider justice from the point of view of the original position would agree upon the same principles of justice generated out of such a thought experiment. Any one person would reach the same conclusion as any other person concerning the most basic principles that must regulate a just society.

The principles that persons in the Original Position, behind the Veil of Ignorance, would choose to regulate a society at the most basic level (that is, prior even to a Constitution) are called by Rawls, aptly enough, the Two Principles of Justice. These two principles determine the distribution of both civil liberties and social and economic goods. The first principle states that each person in a society is to have as much basic liberty as possible, as long as everyone is granted the same liberties. That is, there is to be as much civil liberty as possible as long as these goods are distributed equally. (This would, for example, preclude a scenario under which there was a greater aggregate of civil liberties than under an alternative scenario, but under which such liberties were not distributed equally amongst citizens.) The second principle states that while social and economic inequalities can be just, they must be available to everyone equally (that is, no one is to be on principle denied access to greater economic advantage) and such inequalities must be to the advantage of everyone. This means that economic inequalities are only justified when the least advantaged member of society is nonetheless better off than she would be under alternative arrangements. So, only if a rising tide truly does carry all boats upward, can economic inequalities be allowed for in a just society. The method of the original position supports this second principle, referred to as the Difference Principle, because when we are behind the veil of ignorance, and therefore do not know what our situation in society will be once the veil of ignorance is lifted, we will only accept principles that will be to our advantage even if we end up in the least advantaged position in society.

These two principles are related to each other by a specific order. The first principle, distributing civil liberties as widely as possible consistent with equality, is prior to

the second principle, which distributes social and economic goods. In other words, we cannot decide to forgo some of our civil liberties in favor of greater economic advantage. Rather, we must satisfy the demands of the first principle, before we move on to the second. From Rawls' point of view, this serial ordering of the principles expresses a basic rational preference for certain kinds of goods, i.e., those embodied in civil liberties, over other kinds of goods, i.e., economic advantage.

Having argued that any rational person inhabiting the original position and placing him or herself behind the veil of ignorance can discover the two principles of justice, Rawls has constructed what is perhaps the most abstract version of a social contract theory. It is highly abstract because rather than demonstrating that we would or even have signed to a contract to establish society, it instead shows us what we must be willing to accept as rational persons in order to be constrained by justice and therefore capable of living in a well ordered society. The principles of justice are more fundamental than the social contract as it has traditionally been conceived. Rather, the principles of justice constrain that contract, and set out the limits of how we can construct society in the first place. If we consider, for example, a constitution as the concrete expression of the social contract, Rawls' two principles of justice delineate what such a constitution can and cannot require of us. Rawls' theory of justice constitutes, then, the Kantian limits upon the forms of political and social organization that are permissible within a just society.

b. David Gauthier

In his 1986 book, *Morals by Agreement*, David Gauthier set out to renew Hobbesian moral and political philosophy. In that book, he makes a strong argument that Hobbes was right: we can understand both politics and morality as founded upon an agreement between exclusively self-interested yet rational persons. He improves upon Hobbes' argument, however, by showing that we can establish morality without the external enforcement mechanism of the Sovereign. Hobbes argued that men's passions were so strong as to make cooperation between them always in danger of breaking down, and thus that a Sovereign was necessary to force compliance. Gauthier, however, believes that rationality alone convinces persons not only to agree to cooperate, but to stick to their agreements as well.

We should understand ourselves as individual Robinson Crusoes, each living on our own island, lucky or unlucky in terms of our talents and the natural provisions of

our islands, but able to enter into negotiations and deals with one another to trade goods and services with one another. Entering into such agreements is to our own advantage, and so rationality convinces us to both make such agreements and stick to them as well.

Gauthier has an advantage over Hobbes when it comes to developing the argument that cooperation between purely self-interested agents is possible. He has access to rational choice theory and its sophisticated methodology for showing how such cooperation can arise. In particular, he appeals to the model of the Prisoner's Dilemma to show that self-interest can be consistent with acting cooperatively. (There is a reasonable argument to be made that we can find in Hobbes a primitive version of the problem of the Prisoner's Dilemma.)

According to the story of the Prisoner's Dilemma, two people have been brought in for questioning, conducted separately, about a crime they are suspected to have committed. The police have solid evidence of a lesser crime that they committed, but need confessions in order to convict them on more serious charges. Each prisoner is told that if she cooperates with the police by informing on the other prisoner, then she will be rewarded by receiving a relatively light sentence of one year in prison, whereas her cohort will go to prison for ten years. If they both remain silent, then there will be no such rewards, and they can each expect to receive moderate sentences of two years. And if they both cooperate with police by informing on each other, then the police will have enough to send each to prison for five years. The dilemma then is this: in order to serve her own interests as well as possible, each prisoner reasons that no matter what the other does she is better off cooperating with the police by confessing. Each reasons: "If she confesses, then I should confess, thereby being sentenced to five years instead of ten. And if she does not confess, then I should confess, thereby being sentenced to one year instead of two. So, no matter what she does, I should confess." The problem is that when each reason this way, they each confess, and each goes to prison for five years. However, had they each remained silent, thereby cooperating with each other rather than with the police, they would have spent only two years in prison.

According to Gauthier, the important lesson of the Prisoner's Dilemma is that when one is engaged in interaction such that others' actions can affect one's own interests, and vice versa, one does better if one acts cooperatively. By acting to further the interests of the other, one serves one's own interests as well. We should, therefore, insofar as we are rational, develop within ourselves the

dispositions to constrain ourselves when interacting with others. We should become “constrained maximizers” (CMs) rather than remain the “straightforward maximizers” (SMs) that we would be in a State of Nature (167).

Both SMs and CMs are exclusively self-interested and rational, but they differ with regard to whether they take into account only strategies, or both the strategies and utilities, of those with whom they interact. To take into account the others’ strategies is to act in accordance with how you expect the others will act. To take into account their utilities is to consider how they will fare as a result of your action and to allow that to affect your own actions. Both SMs and CMs take into account the strategies of the other with whom they interact. But whereas SMs do not take into account the utilities of those with whom they interact, CMs do. And, whereas CMs are afforded the benefits of cooperation with others, SMs are denied such advantage. According to Gauthier, when interacting in Prisoner’s Dilemma-like situations, where the actions of others can affect one’s own outcome, and vice versa, rationality shows that one’s own interest is best pursued by being cooperative, and therefore agents rationally dispose themselves to constrain the maximization of their own utility by adopting principles of morality. According to Gauthier, rationality is a force strong enough to give persons internal reasons to cooperate. They do not, therefore, need Hobbes’ Sovereign with absolute authority to sustain their cooperation. The enforcement mechanism has been internalized. “Morals by agreement” are therefore created out of the rationality of exclusively self-interested agents.

4. Contemporary Critiques of Social Contract Theory

Given the longstanding and widespread influence that social contract theory has had, it comes as no surprise that it is also the object of many critiques from a variety of philosophical perspectives. Feminists and race-conscious philosophers, in particular, have made important arguments concerning the substance and viability of social contract theory.

a. Feminist Arguments

For the most part, feminism resists any simple or universal definition. In general though, feminists take women’s experiences seriously, as well as the impact that theories and practices have for women’s lives. Given the pervasive influence of contract theory on social, political, and moral philosophy, then, it is not surprising

that feminists should have a great deal to say about whether contract theory is adequate or appropriate from the point of view of taking women seriously. To survey all of the feminist responses to social contract theory would carry us well beyond the boundaries of the present article. I will concentrate therefore on just three of those arguments: Carole Pateman's argument about the relation between the contract and women's subordination to men, feminist arguments concerning the nature of the liberal individual, and the care argument.

i. The Sexual Contract

Carole Pateman's 1988 book, *The Sexual Contract*, argues that lying beneath the myth of the idealized contract, as described by Hobbes, Locke, and Rousseau, is a more fundamental contract concerning men's relationship to women. Contract theory represents itself as being opposed to patriarchy and patriarchal right. (Locke's social contract, for example, is set by him in stark contrast to the work of Robert Filmer who argued in favor of patriarchal power.) Yet the "original pact" (2) that precedes the social contract entered into by equals is the agreement by men to dominate and control women. This 'original pact' is made by brothers, literally or metaphorically, who, after overthrowing the rule of the father, then agree to share their domination of the women who were previously under the exclusive control of one man, the father. The change from "classical patriarchy" (24) to modern patriarchy is a shift, then, in who has power over women. It is not, however, a fundamental change in whether women are dominated by men. Men's relationships of power to one another change, but women's relationship to men's power does not. Modern patriarchy is characterized by a contractual relationship between men, and part of that contract involves power over women. This fact, that one form of patriarchy was not overthrown completely, but rather was replaced with a different form, in which male power was distributed amongst more men, rather than held by one man, is illustrated by Freud's story of the genesis of civilization. According to that story, a band of brothers, lorded over by a father who maintained exclusive sexual access to the women of the tribe, kill the father, and then establish a contract among themselves to be equal and to share the women. This is the story, whether we understand Freud's tale to be historically accurate or not, of modern patriarchy and its deep dependence on contract as the means by which men control and dominate women.

Patriarchal control of women is found in at least three paradigmatic contemporary contracts: the marriage contract, the prostitution contract, and the contract for

surrogate motherhood. Each of these contracts is concerned with men's control of women, or a particular man's control of a particular woman generalized. According to the terms of the marriage contract, in most states in the U.S., a husband is accorded the right to sexual access, prohibiting the legal category of marital rape. Prostitution is a case in point of Pateman's claim that modern patriarchy requires equal access by men to women, in particular sexual access, access to their bodies. And surrogate motherhood can be understood as more of the same, although in terms of access to women's reproductive capacities. All these examples demonstrate that contract is the means by which women are dominated and controlled. Contract is not the path to freedom and equality. Rather, it is one means, perhaps the most fundamental means, by which patriarchy is upheld.

ii. The Nature of the Liberal Individual

Following Pateman's argument, a number of feminists have also called into question the very nature of the person at the heart of contract theory. The Liberal Individual, the contractor, is represented by the Hobbesian man, Locke's proprietor, Rousseau's "Noble Savage," Rawls's person in the original position, and Gauthier's Robinson Crusoe. The liberal individual is purported to be universal: raceless, sexless, classless, disembodied, and is taken to represent an abstract, generalized model of humanity writ large. Many philosophers have argued, however, that when we look more closely at the characteristics of the liberal individual, what we find is not a representation of universal humanity, but a historically located, specific type of person. C.B. Macpherson, for example, has argued that Hobbesian man is, in particular, a bourgeois man, with the characteristics we would expect of a person during the nascent capitalism that characterized early modern Europe. Feminists have also argued that the liberal individual is a particular, historical, and embodied person. (As have race-conscious philosophers, such as Charles Mills, to be discussed below.) More specifically, they have argued that the person at the heart of liberal theory, and the social contract, is gendered. Christine Di Stefano, in her 1991 book *Configurations of Masculinity*, shows that a number of historically important modern philosophers can be understood to develop their theories from within the perspective of masculinity, as conceived of in the modern period. She argues that Hobbes's conception of the liberal individual, which laid the groundwork for the dominant modern conception of the person, is particularly masculine in that it is conceived as atomistic and solitary and as not owing any of its qualities, or even its very existence, to any other person, in particular its mother. Hobbes's human, is therefore, radically individual,

in a way that is specifically owing to the character of modern masculinity. Virginia Held, in her 1993 book, *Feminist Morality*, argues that social contract theory implicitly relies on a conception of the person that can be best described as “economic man.” “Economic man” is concerned first and foremost to maximize his own, individually considered interests, and he enters into contracts as a means by which to achieve this end. “Economic man”, however, fails to represent all persons in all times and places. In particular, it fails to adequately represent children and those who provide them with the care they require, who have historically been women. The model of “economic man” cannot, therefore, fairly claim to be a general representation of all persons. Similarly, Annette Baier argues that Gauthier’s conception of the liberal individual who enters into the social contract as a means by which to maximize his own individually considered interests is gendered in that it does not take seriously the position of either children or the women who most usually are responsible for caring for those children.

iii. Arguing from Care

Theorizing from within the emerging tradition of care ethics, feminist philosophers such as Baier and Held argue that social contract theory fails as an adequate account of our moral or political obligations. Social contract theory, in general, only goes so far as to delineate our rights and obligations. But this may not be enough to adequately reveal the full extent of what it means to be a moral person, and how fully to respond to others with whom one interacts through relations of dependence. Baier argues that Gauthier, who conceives of affective bonds between persons as non-essential and voluntary, therefore fails to represent the fullness of human psychology and motivations. She argues that this therefore leads to a crucial flaw in social contract theory. Liberal moral theory is in fact parasitic upon the very relations between persons from which it seeks to liberate us. While Gauthier argues that we are freer the more that we can see affective relations as voluntary, we must nonetheless, in the first place, be in such relationships (e.g., the mother-child relationship) in order to develop the very capacities and qualities lauded by liberal theory. Certain kinds of relationships of dependence, in other words, are necessary in the first place if we are to become the very kinds of persons who are capable of entering into contracts and agreements. In a similar vein, Held has argued that the model of “economic man” fails to capture much of what constitutes meaningful moral relations between people. Understanding human relations in purely contractual terms constitutes, according to her argument “an impoverished view of human aspiration” (194). She therefore suggests that we

consider other models of human relationships when looking for insight into morality. In particular, she offers up the paradigm of the mother-child relationship to at least supplement the model of individual self-interested agents negotiating with one another through contracts. Such a model is more likely to match up with many of the moral experiences of most people, especially women.

Feminist critiques of the contractarian approaches to our collective moral and political lives continue to reverberate through social and political philosophy. One such critique, that of Carole Pateman, has influenced philosophers writing outside of feminist traditions.

b. Race-Conscious Argument

Charles Mills' 1997 book, *The Racial Contract*, is a critique not only of the history of Western political thought, institutions, and practices, but, more specifically, of the history of social contract theory. It is inspired by Carole Pateman's *The Sexual Contract*, and seeks to show that non-whites have a similar relationship to the social contract as do women. As such, it also calls into question the supposed universality of the liberal individual who is the agent of contract theory.

Mills' central argument is that there exists a 'racial contract' that is even more fundamental to Western society than the social contract. This racial contract determines in the first place who counts as full moral and political persons, and therefore sets the parameters of who can 'contract in' to the freedom and equality that the social contract promises. Some persons, in particular white men, are full persons according to the racial contract. As such they are accorded the right to enter into the social contract, and into particular legal contracts. They are seen as fully human and therefore as deserving of equality and freedom. Their status as full persons accords them greater social power. In particular, it accords them the power to make contracts, to be the subjects of the contract, whereas other persons are denied such privilege and are relegated to the status of objects of contracts.

This racial contract is to some extent a meta-contract, which determines the bounds of personhood and parameters of inclusion and exclusion in all the other contracts that come after it. It manifests itself both formally and informally. It is an agreement, originally among European men in the beginning of the modern period, to identify themselves as 'white' and therefore as fully human, and to identify all others, in particular the natives with whom they were beginning to come into

contact, as 'other': non-white and therefore not fully human. So, race is not just a social construct, as others have argued, it is more especially a political construct, created to serve a particular political end, and the political purposes of a specific group. The contract allows some persons to treat other persons, as well as the lands they inhabit, as resources to be exploited. The enslavement of millions of Africans and the appropriation of the Americas from those who inhabited them, are examples of this racial contract at work in history (such as Locke's claim that Native Americans did not own the land they lived on because they did not farm it and therefore did not own it). This contract is not hypothetical, as Hobbes describes the one argued for in his *Leviathan*. This is an actual contract, or series of contracts, made by real men of history. It is found in such documents as Papal Bulls and Locke's writings on Native Americans, and acted upon in such historical events as the voyages of discovery made by Europeans and the colonization of Africa, Asia, and the Americas. The racial contract makes possible and justifies some people, in virtue of their alleged superiority, exploiting the peoples, lands, and resources of other races.

From Mills' perspective then, racism is not just an unhappy accident of Western democratic and political ideals. It is not the case that we have a political system that was perfectly conceived and unfortunately imperfectly applied. One of the reasons that we continue to think that the problem of race in the West is relatively superficial, that it does not go all the way down, is the hold that the idealized social contract has on our imagination. We continue to believe, according to Mills, in the myths that social contract theory tells us – that everyone is equal, that all will be treated the same before the law, that the Founding Fathers were committed to equality and freedom for all persons, etc. One of the very purposes of social contract theory, then, is to keep hidden from view the true political reality – some persons will be accorded the rights and freedoms of full persons, and the rest will be treated as sub-persons. The racial contract informs the very structure of our political systems, and lays the basis for the continuing racial oppression of non-whites. We cannot respond to it, therefore, by simply adding more non-whites into the mix of our political institutions, representation, and so on. Rather, we must reexamine our politics in general, from the point of view of the racial contract, and start from where we are, with full knowledge of how our society has been informed by the systematic exclusion of some persons from the realm of politics and contract. This "naturalized" feature of the racial contract, meaning that it tells a story about who we actually are and what is included in our history, is better, according to Mills, because it holds the promise of making it possible for us to

someday actually live up to the norms and values that are at the heart of the Western political traditions.

5. Conclusion

Virginia Held has argued that “Contemporary Western society is in the grip of contractual thinking” (193). Contractual models have come to inform a vast variety of relations and interaction between persons, from students and their teachers, to authors and their readers. Given this, it would be difficult to overestimate the effect that social contract theory has had, both within philosophy, and on the wider culture. Social contract theory is undoubtedly with us for the foreseeable future. But so too are the critiques of such theory, which will continue to compel us to think and rethink the nature of both ourselves and our relations with one another.

6. References and Further Reading

- Baier, Annette. 1988. “Pilgrim’s Progress: Review of David Gauthier, *Morals by Agreement*.” *Canadian Journal of Philosophy* Vol. 18, No. 2. (June 1988): 315-330.
- Baier, Annette. 1994. *Moral Prejudices: Essays on Ethics*. Cambridge: Harvard University Press.
- Braybrooke, David. 1976. “The Insoluble Problem of the Social Contract.” *Dialogue* Vol. XV, No. 1: 3-37.
- DiStefano, Christine. 1991. *Configurations of Masculinity: A Feminist Perspective on Modern Political Theory*. Ithaca, NY: Cornell University Press.
- Filmer, Robert. *‘Patriarcha’ and Other Writings*. Cambridge University Press (1991).
- Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.
- Gauthier, David. 1988. “Hobbes’s Social Contract.” *Noûs* 22: 71-82.
- Gauthier, David. 1990. *Moral Dealing: Contract, Ethics, and Reason*. Cornell: Cornell University Press.
- Gauthier, David. 1991. “Why Contractarianism?” in Vallentyne 1991: 13-30.
- Gilligan, Carol. 1982. *In a Different Voice: Psychological Theory and Women’s Development*. Cambridge: Harvard University Press.
- Hampton, Jean. 1986. *Hobbes and the Social Contract Tradition*. Cambridge: Cambridge University Press.

- Hampton, Jean. 1993. "Feminist Contractarianism." In Antony, Louise M. and Witt, Charlotte (Editors). 1993. *A Mind of One's Own: Essays on Reason and Objectivity*. Boulder CO: Westview Press, Inc.: 1993: 227-255.
- Held, Virginia. 1977. "Rationality and Reasonable Cooperation." *Social Research* (Winter 1977): 708-744.
- Held, Virginia. 1993. *Feminist Morality: Transforming Culture, Society, and Politics*. Chicago: The University of Chicago Press.
- Hobbes, Thomas. 1651a. *Leviathan*. C.B Macpherson (Editor). London: Penguin Books (1985)
- Kavka, Gregory S. 1986. *Hobbesian Moral and Political Theory*. Princeton: Princeton University Press.
- Locke, John. *Two Treatises of Government and A Letter Concerning Toleration*. Yale University Press (2003).
- Macpherson, C.B. 1973. *Democratic Theory: Essays in Retrieval*. Oxford: Clarendon Press.
- Mills, Charles. 1997. *The Racial Contract*. Cornell University Press.
- Nozick, Robert. 1974. *Anarchy, State and Utopia*. New York: Basic Books.
- Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
- Pateman, Carole. 1988. *The Sexual Contract*. Stanford: Stanford University Press.
- Plato. *Five Dialogues*. (Trans. G.M.A. Grube) Hackett Publishing Company (1981).
- Plato. *Republic*. (Trans. G.M.A. Grube, Revised by C.D.C. Reeve) Hackett Publishing Company (1992)
- Poundstone, William. 1992. *Prisoner's Dilemma: John Von Neumann, Game Theory, and the Puzzle of the Bomb*. New York: Doubleday.
- Rawls, John. 1971. *A Theory of Justice*. Harvard University Press.
- Rawls, John. 1993. *Political Liberalism*. Columbia University Press.
- Rousseau, Jean-Jacques. *The Basic Political Writings*. (Trans. Donald A. Cress) Hackett Publishing Company (1987).
- Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.
- Vallentyne, Peter. (Editor). 1991. *Contractarianism and Rational Choice: Essays on David Gauthier's Morals by Agreement*. New York: Cambridge University Press.

Author Information

Celeste
Email:
Hamilton
U. S. A.

Friend
cfriend@hamilton.edu
College