CRITICAL NOTICE

Hobbes and the Social Contract Tradition. Jean Hampton. New York: Cambridge University Press, 1986. Reviewed by Theodore S. Zenzinger. University of Kansas.

The central claims of Hobbes's *Leviathan* are well known. Humans, to fulfill their desires, including a primary desire for self-preservation, attempt to secure and wield as much power as possible. Since all possess by nature the liberty to use power as they see fit, and since all are roughly equal in strength of mind and body, it can be expected that there will be competition for power, which will cause quarrels, diffidence, and vain glory. The combination of these transforms the state of nature into the state of war where life is nasty, brutish and short. Since the root cause of the war of all against all is the natural liberty of each to use power as s/he sees fit, reason suggests that individuals contract with one another to surrender this liberty to an absolute sovereign.

Traditionally, two questions have been asked of Hobbes's argument. First, if humans are rational, why do they not see the benefits of cooperation, temper their actions and so avoid the war of all against all? Second, if humans are not able to cooperate with one another without the presence of an absolute sovereign, how can they ever hope to contract with one another to institute an absolute sovereign? Jean Hampton in her *Hobbes and the Social Contract Tradition* attempts to give a reading of Hobbes that answers these two important questions. While I do not think Hampton's reading of Hobbes avoids the problems latent in the two questions posed above, I do believe that she offers the strongest presentation of Hobbes's argument to date; it will therefore be instructive to determine exactly where her attempt encounters difficulties.

Hampton begins her reading of Hobbes by explaining how rational individuals fail to cooperate in the state of nature. The key to her explanation is the short-sightedness of individuals:

> The account would contend that many people fail to appreciate the long-term benefits of cooperation and opt instead for the short-term benefits of non-cooperation, and the rest are legitimately fearful enough of this shortsightedness afflicting their partners to doubt that cooperation would have any educative effect. [Hampton 1986, p. 81]

An important assumption here is that there are two types of individuals: the short-sighted and the non-short-sighted (hereafter called the prudent). An additional assumption is that there are long-term benefits of cooperation; that over a long period those who cooperate with their fellows

can expect to do better than those who do not. Hampton rests this assumption on the work of Robert Axelrod[1] who argues that in some situations individuals who chose cooperative strategies did better than those who did not. Axelrod carefully noted, however, that cooperative strategies did worse in some environments than non-cooperative strategies. Thus, it cannot be expected that there will always be long term benefits to cooperation.[2] An additional assumption is that the short-sighted are not educable; they possess a fixed disposition for short-sightedness.

Two types of short-sightedness may afflict individuals.

> First, the iterated nature of the PD game might not be obvious to some of the less intellectually talented inhabitants of this state, who either never realize that they should take a multiplay perspective or are not mentally acute enough to work out the long-term benefits of cooperation. [Hampton 1986, p. 82][3]

---

[1] Robert Axelrod, *The Evolution of Cooperation*. New York: Basic Books, 1984.

[2] An example might help to clarify this point. Suppose that the world consists of four individuals, three of whom have adopted the strategy "Always Defect," and one who adopted the strategy "Tit for Tat" (cooperate on the first move and then do what the other individual did on the previous move). The three who chose "Always Defect" will do better than the one who chose "Tit for Tat," for each of the three will be able to take advantage of the "Tit for Tat" player before that individual responds with defection. Since the "Tit for Tat" player can never get the others to cooperate and can never take advantage of them, s/he will never be able to make up for being taken advantage of on the first move. Someone who chose "Always Cooperate" in the same environment would fair even worse.

[3] Hampton's model for interaction in the state of nature is the Prisoner's Dilemma. The Prisoner's Dilemma is usually represented by a matrix of the form:

		B	
		Cooperate	Not Cooperate
	Cooperate	(2,2)	4,1)
A			
	Not Cooperate	(1,4)	3,3)

Payoffs here are given in rankings with 1 indicating the outcome most preferred by an individual (by convention, A's preferences are given first), and so on. No matter what B chooses, A does better by choosing not to cooperate, for 1 > 2 and 3 > 4. The same holds true for B. When Hampton

Let us call individuals who fail to recognize that they are in an iterated Prisoner's Dilemma the "cognitively short-sighted."[4]

> Second, and more important, even if it did occur to them that they might be in an iterated PD game in which substantial long-term benefits make cooperation rational, they might decide, nonetheless, to reject such a characterization of their situation on the grounds that the prospect of future contractual interactions with their present partners is too remote to warrant it. [Hampton 1986, p. 82]

Let us call those who so heavily discount the future as to turn the iterated Prisoner's Dilemma into a single play Prisoner's Dilemma the "affectively short-sighted."

Both types of short-sighted individuals treat encounters in the state of nature as single play Prisoner's Dilemmas, and in such situations, the rational action is not to cooperate.[5] If short-sighted individuals comprise the entire human population, the explanation for the origin of conflict would be straightforward. Hampton notes, however, that assuming the entire population consists of short-sighted individuals is unreasonable and that conflict will arise even if the population has both short-sighted and prudent individuals.[6]

Even if the prudent recognize that humans in the state of nature are in an iterated Prisoner's Dilemma, the presence of the short-sighted forces them to choose not to cooperate. For, the prudent are conditional cooperators, cooperating only if they expect their partners to do so as well. If the prudent have no knowledge of the dispositions of their partners, and if the cost of exploitation is very high, then the prudent will be forced to choose non-cooperation out of self-defense. However, if several prudent individuals come to recognize one another as prudent, then there may be scattered pockets of cooperation. Provided that there are not too many pockets of cooperation, the fact that there are such pockets does not imply

---

speaks of iteration, she has in mind the indefinite repetition of this choice situation for the same individuals. For a useful introduction to this problem see Richmond Campbell and Lanning Sowden, eds. *Paradoxes of Rationality and Cooperation*. Vancouver: The University of British Columbia Press, 1985.

[4] I owe this label and the equally useful "affectively short-sighted" to Ann Cudd.

[5] See footnote 4.

[6] See Hampton 1986, p. 81ff.

there is no need for a sovereign (though it may imply that hobbesians do not require the services of an absolute sovereign[7]).

It has just been established that with the odd exception, hobbesians will choose not to cooperate with one another in the state of nature (shortly, we will examine the conditions under which cooperation is rational in greater detail). How, then, can they cooperate (contract) with one another to institute a sovereign? If their situation is that of the Prisoner's Dilemma, then there seems little hope that a sovereign will be instituted.[8] Fortunately, argues Hampton, this is not the situation faced by hobbesians:

> ...the fact that people are continually in (what they are forced to regard as) single-play prisoner's dilemmas in the state of nature does not mean that they regard escaping this PD-prone situation as itself posing a prisoner's dilemma for them. [Hampton 1986, p. 148]

Instead, of taking themselves to be in a Prisoner's Dilemma,

		B	
		Surrender	Do not Surrender
	Surrender	(2,2)	(4,1)
A			
	Do not Surrender	(1,4)	(3,3)

(figure 1)

individuals take themselves to be in a coordination problem,

---

[7] Hampton's book contains an interesting discussion of the problem of the absolute sovereign, which I will not be able to address in detail. The earliest commentators of Hobbes noted the peculiarity that hobbesians are to create an absolute sovereign while retaining a right to self-defense that is virtually boundless. Many have argued that the retention of this right undermines the authority of the absolute sovereign, thereby precluding her/his ability to lead the people out of the state of war. Hampton's response to this complaint is to argue that hobessians do not need an absolute sovereign; that Hobbes recognized this, and provided a description of an agency social contract as a fallback position. While I find Hampton's reading of Hobbes on this point interesting, it does not avoid the problems created by the introduction of the short-sighted.
[8] See footnote 4.

	B	
	Surrender	Do not Surrender
A Surrender	(1,1)	(2,4)
Do not Surrender	(4,2)	(3,3)

(figure 2)

Figure 2 represents a situation where a sovereign has been chosen, and where A and B must decide whether to surrender to her or not. Since both A and B fear the combined might of the sovereign and the other (as indicated by the preferences for outcomes), each will find it rational to surrender to the sovereign. This will be true even for short-sighted individuals.

All Hobbesians are interested in being sovereign: it offers each the best chance to secure her/his life. Yet, only one can be sovereign.[9] Determining who will be sovereign is a "Battle of the Sexes"[10] problem, argues Hampton:

	B	
	A is Sovereign	B is Sovereign
A A is Sovereign	(1,2)	(3,3)
B is Sovereign	(3,3)	(2,1)

(figure 3)

Initially, at least, each individual would prefer to be sovereign; if this is not possible, each would prefer that there were a sovereign rather than multiple sovereigns (in effect, no sovereign). Thus individuals find themselves in the situation depicted by figure 3.

The problem is that there are two equilibria, (1,2) and (2,1). In order for individuals to coordinate their actions, they must be able to decide which of the two equilibria outcomes will be their goal. Hampton outlines several ways to accomplish this.

First, one individual may be in a position to make side payments to the other e.g., A might promise B that the next time they are in a similar situation she will yield to his preferences. This method of breaking ties is not available to the short-sighted, however. The cognitively short-sighted will not perceive the advantage to be gained by making or receiving side payments since they are unable to look ahead to future encounters. The

---

[9] Hobbes did allow that the sovereign be a body of people, but he believed this to be inferior to the sovereignty of one.
[10] For a description of this game see R. Duncan Luce and Howard Raiffa, *Games and Decisions.* Mineola: Dover Publications, 1989 (reprint).

affectively short-sighted are not interested in the future and so will not be persuaded by the promise of future reward.

Second, A and B might negotiate and by this process make one equilibrium salient by agreement. Setting aside the problem of choosing a bargaining theory, can short-sighted individuals perform the required negotiations? I do not believe that they can. Both individuals will have two options in each round of negotiations: to give in, or to hold out. The negotiations will be successful when one of them decides to give in, but it is unlikely that this will occur. The cognitively short-sighted will never be able to agree; they will always choose to hold out for they will not perceive that they are involved in anything other than a single-play game. The affectively short-sighted will not be interested in the future benefit to be gained by a negotiated outcome. It is likely that they will quickly loose interest in negotiation and walk away.

Hampton recognizes that the cost of negotiation may be high; thus she suggests that individuals might instead choose a sovereign by successive balloting.[11] In the first round, each individual will presumably vote for himself. In subsequent rounds, however, individuals will perform expected utility calculations, and shift their allegiance from themselves to the individual who is most likely to win. Eventually, one individual will receive most votes cast. The difficulty with choosing a sovereign by successive balloting is that the ability to perform expected utility calculations is a trait of the prudent. The cognitively short-sighted cannot be expected to change their vote, for they will not perceive that they are involved in a continuing process. The affectively short-sighted will not be interested in anything but the current round of voting, and thus are likely to alter their initial vote. It thus seems that the process of selecting a sovereign involves only the prudent.

One final problem remains to be solved, that of empowerment. In order for the sovereign to enforce her orders she must have some instrument with which to enforce her will. For example, the sovereign needs individuals who are willing to capture dangerous criminals, police the streets, et cetera. All prefer that desperate criminals be apprehended, yet all also prefer not to face the risk associated with participating in their capture. Furthermore, each knows that even if he does not participate in the capture of the criminal provided that enough others do so, he will continue to share in the benefits accrued by having a dangerous criminal in jail. Since all share this preference and possess this knowledge, all are unlikely to willingly become members of the sovereign's punishment cadre. This problem is often called "The Free Rider Problem."

---

[11] If a large number of voting rounds is required, it is difficult to see how voting is any less costly than negotiation.

If the Free Rider Problem is correctly modelled by the Prisoner's Dilemma[12] then it is unlikely that the sovereign will have a punishment cadre, and being unable to enforce her orders, will not be able to ensure that contracts are kept and the general peace secured. Fortunately, argues Hampton, the Free Rider Problem is not simply an n-person Prisoner's Dilemma. Hampton uses a careful examination of Hume's meadow draining example to make her case. Her key claim is that while Prisoner's Dilemmas have dominant actions, free-rider problems do not. This can be illustrated by a simplified version of Hume's meadow draining example. Consider the case of three individuals whose property adjoins a boggy meadow. All would benefit if the meadow were drained. If we suppose that only two need work together to drain the meadow, we can represent the choice situation facing one of them, A, thusly:

		Others		
		N-2 Help	N-1 Help	N-Help
	Help	(5)	(3)	(2)
A				
	Not Help	(4)	(4)	(1)

(figure 4)[13]

Inspection reveals that there is no dominant strategy; individuals do not prefer no helping to helping under all circumstances. The key cases will be those where n-1 of the persons needed to drain the meadow have agreed to do so. In these cases individuals my prefer helping to not helping.

Putting aside the issue of whether Hampton correctly models the free-rider problem, reveals for us what is the important question. Will short-sighted individuals choose to help when placed in the game depicted in figure 4? Hampton suggests that the game is a three way battle of the sexes problem. As I argued earlier, however, it is unlikely either the cognitively short-sighted or the affectively short-sighted have the requisite talents to cooperate in such games.

---

[12] See footnote 4 for the PD matrix. If "cooperate" is replaced with "Become a Cop," and "Not Cooperate" is replaced with "Don't Become a Cop," it is easy to see that individuals will not choose to become cops (join the punishment cadre).

[13] Payoffs here are given in rankings with 1 indicating an individual's most preferred outcome. It is assumed that individuals most prefer the outcome where the meadow is drained but they do not have to help, and least prefer that outcome where they attempt to drain the meadow with no assistance. See Hampton 1986, p. 178.

It would be tempting to conclude that Hobbesians are not able to choose a sovereign, or empower a sovereign, and thus that they are forever trapped in the war of all against all. But this would be premature, for Hampton would rightfully remind us that we have failed to consider the presence of the prudent. At first blush appeal to the prudent might seem to save Hampton's account. The prudent will have no problem developing a voting strategy, and will successfully empower a sovereign. Yet, Hampton cannot allow the entire population of the state of nature to be prudent, if it were, there would be no conflict and thus no need for a sovereign. What Hampton must do is to establish that there is some ratio of short-sighted to prudent such that there will be sufficient conflict to warrant the institution of a sovereign, yet sufficient cooperation to institute successfully the needed sovereign.

Let us begin afresh by looking at the conditions under which the prudent will select non-cooperation over cooperation when in the state of nature. For illustrative purposes let us examine the encounter between two individuals in the state of nature. A, who is prudent (which I will indicate hereafter by "Ap"), and B who disposition is unknown to A. Ap wishes to discover whether the conditional expected utility (CEU) of cooperating (hereafter indicated by "C") is greater or less than the CEU of non-cooperation (hereafter indicated by "-C"). Ap, then, is interested in whether the CEU (C), which equals p(B [C]/Ap [C])u1 + p(B [-C]/Ap [C])u2, is less or greater than the CEU (-C), which is equal to p(B [C]/Ap [-C])v1 + p(B [-C]/Ap [-C])v2.[14] Where:

u1 = the cardinal utility to A if she and B cooperate
u2 = the cardinal utility to A if she cooperates and B does not
v1 = the cardinal utility to A if she does not cooperate and B does
v2 = the cardinal utility to A if she and B do not cooperate

---

[14] In words, we are interested in whether the conditional expected utility to A of choosing cooperation, which is equal to the probability that B will choose cooperation given that A chooses cooperation times the utility to A if both she and B choose cooperation, plus the probability that B will choose not to cooperate given that A chooses to cooperate times the utility to A if she cooperates and B does not, is greater or less than the conditional expected utility to A if choosing not to cooperate, which is equal to the probability that B chooses cooperation given that A chooses not to cooperate times the utility to A if she does not cooperate and B does cooperate, plus the probability that B chooses not to cooperate given that A chooses not to cooperate times the utility to A if neither she nor B cooperates. For a useful discussion of conditional probabilities, see *Paradoxes of Rationality and Cooperation* referenced in footnote 4.

Being prudent, Ap is disposed to cooperate when others cooperate with her, and not to cooperate when others do likewise. In addition Ap believes that the decision she and B make will be symmetrical; that the same factors that lead her to choose cooperation or non-cooperation will lead B to choose cooperation or non-cooperation. Thus, Ap is interested in comparing p(B [C]/Ap [C])u1 with p(B [-C]/Ap [-C])v2. When p(B [-C]/Ap [-C])v2 > p(B [C]/Ap [C])u1, Ap will choose not to cooperate.

Since Ap bases her choice on what she expects B to do, we must find the circumstances under which B will cooperate and those under which he will not. If B is prudent, he will cooperate with Ap when he successfully recognizes her as prudent. Thus, p(B [C]) = p(Bp and Br) = p(Bp)p(Br) = p(Bpr).[15] (Here, "r" stands for recognizing the disposition of the other and "-r" will indicate a failure to recognize the disposition of the other.) B will not cooperate with Ap when either he is short-sighted (which I will write "Bs") or when he fails to recognize Ap as prudent. Thus, p(B [-C]) = p(Bs or B-r) = p(Bs) + p(B-r). Making the appropriate replacements into p(B [-C]/Ap [-C])v2 and p(B [C]/Ap [C])u1, Ap is interested in comparing p(p(Br)/Ap[C])u1 with p(p(Bs) + p(B-r)/Ap [-C])v2. In making this comparison we can remove from comparison what Ap would do by alternately setting p(Ap [C]) and p(Ap [-C]) at 1. The relevant comparison, then, is between (p(Bp)p(Br))u1 and (p(Bs) + p(B-r))v2.

It quickly becomes apparent that p(Br) plays a critical role in determining the ratio of short-sighted to prudent necessary for Ap to choose non-cooperation over cooperation. Consider tables 1, 2, and 3. In table 1 p(Br) is equal to .5:

p(Bp) x p(Br) x u1 = 1 p(Bs) = p(B-r) x v2 = 1 C -C

.1	.5	3	.15	.9	5	1	1.4		x
.2	.5	3	.30	.8	5	1	1.3		x
.3	.5	3	.45	.7	5	1	1.2		x
.4	.5	3	.60	.6	5	1	1.1		x
.5	.5	3	.75	.5	5	1	1.0		x
.6	.5	3	.90	.4	5	1	.90		
.7	.5	3	1.05	.3	5	1	.80	x	

(table 1)

In table 2 p(Br) is equal to .6:

p(Bp) x p(Br) x u1 = 1 p (Bs) + p(B-r) x v2 = 1 C -C

---

[15] In words, the probability that B will choose to cooperate is equal to the probability that he is prudent and he recognizes A as prudent

								C	-C
.1	.6	3	.18	.9	.4	1	1.3	x	
.2	.6	3	.36	.8	.4	1	1.2	x	
.3	.6	3	.54	.7	.4	1	1.1	x	
.4	.6	3	.72	.6	.4	1	1.0	x	
.5	.6	3	.90	.5	.4	1	.90		
.6	.6	3	1.08	.4	.4	1	.80		x

(table 2)

In table 3 p(Br) is equal to .7:

$$p(Bp) \times p(Br) \times u1 = 1 \quad p(Bs) + p(B\text{-}4) \times v2 = 1 \quad C \quad \text{-}C$$

								C	-C
.1	.7	3	.21	.9	3	1	1.2	x	
.2	.7	3	.42	.8	3	2	1.1	x	
.3	.7	3	.63	.7	3	1	1.0	x	
.4	.7	3	.84	.6	3	1	.9	x	
.5	.7	3	1.05	.4	3	1	.7		x

(table 3)

Note the effect of the increase in p(Br): as it increases the ratio of short-sighted to prudent needed for -C to be chosen increases.

If Hampton can provide a convincing argument at p(Br) is .5 or .6, then her reading of Hobbes is only subject to the difficulty that it provides for a sovereign by conquest. As we noted earlier, the short-sighted are not able to participate in the institution of a sovereign, though once a sovereign is instituted by the prudent, the short-sighted can be forced to surrender their natural liberty. In instances where p(Br) is .5 or .6, one would not expect the short-sighted to present much of a problem, for in these instances there need not be many short-sighted individuals for prudent individuals to choose non-cooperation over cooperation. The relatively large population of prudent individuals should have little trouble controlling the short-sighted. But, where p(Br) is greater than .6, an additional problem emerges for Hampton's reading of Hobbes. As p(Br) increases, prudent individuals will choose non-cooperation only when the short-sighted outnumber the prudent. The prudent will still recognize the value of a sovereign and will act together to institute one, but their ability to compel the short-sighted to join the commonwealth, as well as their ability to control them, will be impaired by their fewer numbers.

There is reason to suspect that p(Br) will be high. First, if it is low (.5 or .6) the large number of prudent individuals probably will form extensive pockets of cooperation, thus lessening the need for an absolute sovereign. In order, then, to have widespread conflict, the number of prudent individuals must be small, but this implies that p(Br) must be high.

Second, and more telling, it should be easy for the prudent to recognize the short-sighted: all they need do is devise a test that reveals the inability of the short-sighted to look beyond immediate gain. Again, this suggests that p(Br) will be high, and thus that the number of prudent individuals will be significantly smaller than the number of short-sighted individuals. This in turn suggests that on Hampton's reading Hobbesians cannot form a stable commonwealth.

To sum up, self-interested individuals whose primary preference is for their reservation, transform the state of nature into the war of all against all. This same self-interest and preference leads them to renege on contracts, thus making the institution of an absolute sovereign by contract impossible. An absolute sovereign is necessary, however, for only an absolute sovereign, who has the power to enforce the laws of nature, allows individuals to escape the war of all against all. Although individuals will not willingly agree to institute a sovereign they can be compelled to do so by force. In a commonwealth by acquisition, the original preferences of individuals are overlain with new, threat induced preferences. While there is a continuous threat, the threat induced preferences will make it rational for individuals to abide by the laws of nature. The ability of the sovereign to project a continuous threat depends, however, on her ability to establish a viable punishment cadre. Yet individuals will not willingly become members of a punishment cadre for by doing so they place their lives at risk. Unless individuals are faced with a viable threat, they will not become members of a punishment cadre. Unfortunately the viable threat that would induce them to become members of a punishment cadre must itself be provided by a punishment cadre. Hobbesians, then, are unable to institute a sovereign by contract or be members of a stable commonwealth created by conquest. The self-interest and preference that doom the state of nature to perpetual war, forestalls the attempt of individuals to escape from it. To escape from the war of all against all, individuals must be able to complete contracts. Could they do this they would need no sovereign to enforce the laws of nature, which compel individuals to keep the contracts they have made. Thus, if Hobbesians need an absolute sovereign, they are unable to institute one; could they institute one, they would have no need of one. In brief, this is Hobbes's dilemma.

In may view Hampton is attempting to steer between the two horns of Hobbes's dilemma. On the one hand, if the population of the state of nature consists only of short-sighted individuals, there will be conflict, but there can be no sovereign instituted by agreement. On the other hand, if the population of the state of nature consists entirely of prudent individuals, an agreement instituting a sovereign becomes possible, but there will be no conflict in the state of nature, and thus no need for a sovereign. Hampton attempts to find a middle ground by arguing that the population of the state of nature has both short-sighted and prudent individuals. If the proper ratio of short-sighted to prudent can be

determined, then the presence of the short-sighted will lead to conflict; and the presence of the prudent will lead to the institution of a sovereign. The problem with this middle ground is that the short-sighted are not party to the agreement that creates the sovereign; they become members of the commonwealth by force of arms, not by voluntary agreement. As unwilling participants in the commonwealth, they can be expected to disobey the sovereign whenever they believe they can get away with it. Their non-compliance can only be contained if the sovereign's punishment cadre is large enough to present a viable threat, but this implies that the ratio of prudent individuals to short-sighted individuals is large. Given a larger ratio of prudent to short-sighted, however, the likelihood of conflict decreases and with it the likelihood that there is need for a sovereign. Thus, the dilemma which Hampton sought to avoid threatens to reappear.

   While I do not believe that Hampton's device of dividing hobbesians into the short-sighted and prudent avoids the problems latent in Hobbes's social contract theory, I did find her book valuable. Hampton's careful presentation of Hobbes provokes thoughtful response; her equally careful and thoughtful attempt to save as much of that theory as possible led me to take a long second look at the promise and limitations of social contract theory.