

1974

Paying the Penalty: The Role of Punishment in Theories of Justified Civil Disobedience

Daniel Michael Farrell

Follow this and additional works at: https://digitalcommons.rockefeller.edu/student_theses_and_dissertations

 Part of the [Life Sciences Commons](#)

PAYING THE PENALTY:
THE ROLE OF PUNISHMENT IN THEORIES
OF JUSTIFIED CIVIL DISOBEDIENCE

A Thesis submitted to the Faculty of The Rockefeller University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

by

Daniel M. ^{Michael} Farrell, B.A.
_{IN}

April 15, 1974
The Rockefeller University
New York

E C

D4711.6

245

2

PREFACE

It is a pleasure to thank all of those who have helped me to complete this project, both with their advice and by their encouragement. It would be impossible to mention everyone to whom I am indebted, but some have made such a great contribution to my progress, such as it is, that I wish to take this opportunity to express my gratitude formally.

I am indebted, first and foremost, to Professor Joel Feinberg. What is rather a worn cliché is nonetheless profoundly true: without his help I would not have gotten this work done. I would also like to thank the many people -- faculty, staff, fellow-students and friends -- who made my years at The Rockefeller University so rewarding, both personally and intellectually. Special thanks are due to Robert Schwartz for reading through the final manuscript and commenting upon it.

A number of philosophers at a number of other institutions have also helped me with their criticisms. I benefited from the responses to papers read at The Ohio State University, The University of Illinois, Chicago Circle, The University of Maryland at Baltimore County, and York University, Toronto.

My friends and students at Ohio State have been extremely generous, both in listening to my ideas and in responding to them. I am especially grateful to James Jeffers, John Josephson, Larry Peterson, and Virginia Tuttle. Many others will excuse me, I hope, if I thank them without naming them.

Several foundations contributed to my support in graduate school, and I thank them here: The Woodrow Wilson Fellowship Foundation, The Danforth Foundation, and The Rockefeller University.

I am deeply indebted to others for help and guidance of a special sort: Fritz Schmidl, Rudolph Ehrensing, Richard Sallick, and Irving Pine. Their counsel has seen me through some difficult times.

Finally, my family has been a constant source of personal comfort and encouragement, and my two best friends, James Jeffers and Virginia Tuttle, have made working, and living the intellectual life, a genuine joy.

I would like to dedicate this work to Professor Joel Feinberg, especially, and to all the teachers who have helped and encouraged me with their advice and by their example.

TABLE OF CONTENTS

PREFACE	ii
CHAPTER I: SOME RECENT ARGUMENTS FOR THE WILLINGNESS REQUIREMENT	1
1. Statement of the problem and strategy of the thesis as a whole	1
2. Some preliminary distinctions	4
3. Bedau: the "definitional" move	5
4. M. Cohen: the "tactical" move	7
5. Brown: paying the penalty as a <u>necessary condition</u> of justified civil disobedience	9
6. Konvitz: paying the penalty as an <u>essential</u> <u>characteristic</u> of justified civil disobedience	14
7. C. Cohen: paying the penalty and the principles of punishment	21
Notes	160
CHAPTER II: RAWLS'S ARGUMENTS FOR THE WILLINGNESS REQUIREMENT	29
1. Definitional and tactical rationalia <u>vs.</u> paying the penalty as a necessary condition of justified civil disobedience: some additional remarks	29
2. Rawls's view of the role of the willingness requirement	33
3. The argument that "publicity" and nonviolence entail the willingness requirement	38
4. The argument(s) from the notion of "fidelity to law"	40
5. Fried's defense of the willingness requirement: the argument based on the principle of fair play	49
Notes	164
CHAPTER III: THE "LEGALIST" TRADITION AND SOCRATES' ARGUMENTS IN THE <u>CRITO</u>	53
1. The varieties of "legalism", historical and modern	53
2. Similarities between the views of Socrates and Rawls	56
3. Preliminary remarks on the <u>Crito</u> and its relevance	59
4. Some assumptions about the argument in the <u>Crito</u> and about its general structure	61
5. Explication and evaluation of the principal arguments in the dialogue	62

6. The paradox involved in Socrates' two main claims	68
Notes	167
CHAPTER IV: RAWLS'S ARGUMENTS FOR OBEYING JUST AND UNJUST LAWS	72
1. Introductory remarks	72
2. The argument from ideal theory	74
3. Problems with this argument	77
4. The argument from nonideal theory	81
5. Problems with this argument	86
6. The argument for obeying unjust laws: the principle of fair play	88
7. The argument for obeying unjust laws: the added empirical premiss	91
8. Explaining and evaluating the new premiss	95
9. Some objections to the argument for obeying unjust laws: general strategy	98
10. Some cases that seem to undermine Rawls's view and the problem of generalization	100
Notes	172
CHAPTER V: THE ARGUMENT BY ANALOGY	107
1. Strategy of this chapter	107
2. The argument by implication	109
3. The argument that justification entails nonpunishment . .	111
4. Preliminary remarks about the argument by analogy	115
5. The principal empirical claim in the argument by analogy	117
6. The formal version of the argument by analogy	121
Notes	174
CHAPTER VI: JUDICIAL IMPLEMENTATION: SOME TENTATIVE SUGGESTIONS	128
1. Statement of the problem	128
2. Dworkin's proposal and objections to it	131
3. Van Dyke's proposal and problems with it	135
4. My own proposal	139
5. Some examples	144
6. Defense of my proposal: theoretical considerations . . .	145
7. Defense of my proposal: practical considerations	151
Notes	177

CHAPTER I

SOME RECENT ARGUMENTS FOR THE WILLINGNESS REQUIREMENT

1. The problem of civil disobedience and its justification has received a great deal of attention in the past decade, and it seems safe to say that a consensus of sorts has been reached with respect to the basic questions of (i) whether such dissent can be morally justified in a legitimate constitutional democracy; and, if it can be justified, (ii) roughly what the conditions are under which this is the case. The received view among liberal political theorists, as I understand it, is that civil disobedience can be morally justified, under certain conditions, even in a more or less just democratic society.¹ What's more, there is a remarkable degree of agreement among contemporary thinkers as to roughly what the conditions are under which civil disobedience is justified.²

There are dissenters to what I call the received view,³ of course, but the degree of agreement among most contemporary writers is quite impressive. My principal aim in the following chapters is to show that the received view with respect to the justification of civil disobedience is radically defective in at least one respect, defective in a way that is of considerable intellectual and practical consequence. For according to virtually every theory of civil disobedience in the recent literature, an essential part of being morally justified in an act of civil disobedience is a willingness to accept the full legal consequences of that act.⁴ This "willingness requirement", as I shall call it, is incorporated in different ways into different theories, but for the most part recent theorists agree that civil disobedience is morally justifiable only, inter alia, if the dissenter is willing to pay the penalty for his disobedience. I shall show that this contention is mistaken and that a willingness to accept the consequences is not required, morally speaking,

for the justification of civil disobedience, even in a more or less just constitutional democracy.

As I have said, and as I shall show in considerable detail in section 3 below, the willingness requirement can be incorporated into theories of civil disobedience in importantly different ways. Thus, some writers claim that an illegal act is, by definition, not "civil disobedience" unless the dissenter is willing to accept the consequences of his conduct. Other theorists, while avoiding this definitional maneuver, have been anxious to show that, tactically speaking, civil disobedience is not likely to be effective unless the disobedients exhibit a willingness to pay the penalty for their disobedience and hence that they ought to do so. I shall have much more to say about these matters below (see section 3), but I want to emphasize from the start that I am not interested in arguments for the willingness requirement that proceed primarily on either definitional or tactical grounds. My concern is with arguments that make a willingness to pay the penalty a matter of moral and political principle. I need to say something about this last point before going on.

When a thinker claims that civil disobedience can be "justified", even in a reasonably just democratic society, for example, he is usually claiming not that disobedience to the law can be legally justified but that despite its obvious (and intentional) illegality such conduct can be morally justified.⁵ And one point which nearly all of these thinkers emphasize is the point I've made my quarry: the claim that to be morally justified in his illegal action the civil disobedient must be willing to accept the legal penalties for an act that is itself morally right and perhaps even morally obligatory.

That in general one should expect to have to suffer for one's moral righteousness, as a sad matter of fact, is perhaps not surprising; but that one should be required, as a matter of moral principle, to expect and accept punishment for doing what is otherwise morally right and perhaps even morally obligatory, is, it seems to me, somewhat paradoxical. And in any case, even if this apparent paradox can be dissipated, if someone wishes to argue that I have an obligation to accept that punish-

ment willingly, when I have rightly refused to obey an unjust law, it seems that the proponent of this view ought to bear the burden of proof. We should expect him to give us, first, a general account of what it is for an action to be morally right (even if plainly illegal) and, second, a very detailed account of why it is that, morally speaking, it is right to punish right conduct (when the latter is illegal).

Curiously, no more than a handful of recent theorists have made any effort to discharge this burden of proof. It is worth noticing why this is so and how it is that the burden of proof is assumed to lie where it is taken to lie in recent discussions of the justification of civil disobedience, and I shall get into these matters below (see Chapters III and IV). It will be my contention that almost all recent discussions of civil disobedience, while anxious to vindicate the right of the individual to disobey unjust laws as a last resort, have taken place in a framework that makes it quite natural to put the entire (moral) burden of proof on the individual in establishing his right to disobey, rather than on society in making good its right to enforce legislation that is at best controversial, from the point of view of justice and morality, and is in some cases clearly unjust and immoral.

But before doing this I wish to glance, in this chapter, at some of the more articulate defenses of the willingness requirement in the recent literature, and suggest roughly where they go wrong. In Chapter II I want to study the best explicit defense I know of for the willingness requirement: Professor John Rawls's arguments in Chapter VI of his recent book, *A THEORY OF JUSTICE*. I shall show that Rawls's (explicit) arguments are really no better than the arguments of his predecessors.

In Chapters III, IV, and V I shall elaborate what seems to me a far more impressive defense of the willingness requirement, a defense that is suggested by some recent work of Rawls on a related topic. This defense fails too, however, and in Chapter VI I shall show why this is so.

My aim in Chapter VI is to show how the willingness requirement gains whatever plausibility it has from our tendency to think that it would be impossible -- or disastrous -- to allow legal defenses for civil

disobedience, defenses that would be based on the claim that the illegal act was morally justified. In this last chapter I shall attempt some crude legal draftsmanship and suggest roughly how defenses based on the moral propriety of disobedience to the law might be worked into our judicial system. A great deal of work obviously remains to be done in this last area. My principal aim is to show why we must face the task of doing it and to suggest that this is not an impossible task.

2. I indicated above that what I call the "willingness requirement" is built into different theories of civil disobedience in different ways. It would be nice if this were not so. For then I could simply say "Here's the view I'm after," point to it in a number of recent theories, and show how none of the arguments in its favor does the job required of it. My task is complicated, however, by the fact just mentioned: not every theory I wish to discuss straightforwardly contends that civil disobedience is simply not justified unless the disobedient is ready and willing to accept all the legal consequences of his illegal act. In some cases we shall indeed find contentions as straightforward as this one. But in the most interesting cases, we shall not find that the willingness requirement is made an explicit necessary condition of justified civil disobedience. Part of our task, then, will be to determine just what role paying the penalty plays in a given theory of civil disobedience.

The following distinctions will enable me to isolate the precise component of recent liberal political theory that I am interested in. There are, to begin with, at least three very different sorts of reasons that are frequently given for saying that anyone contemplating civil disobedience ought to be willing to pay the penalty for his illegal act. The first is simply that he will not be perpetrating "civil disobedience" if he does not express such a willingness. The second is that he is not likely to be successful in his undertaking if he fails to accept the legal consequences willingly. The third is that he cannot be justified in what he is doing unless he is willing to pay the penalty for what he does.

The first move is what I call a "definitional" one. The second is

obviously a "tactical" one. And I shall say that arguments involving considerations of the third sort above are arguments that make paying the penalty a matter of principle: the act is not morally justified if the relevant willingness is not present.

Unfortunately, none of the theories we shall be studying arranges things quite so tidily as this. However, I think that we can tidy them up without distorting them. That is, I think we can separate those thinkers who defend the willingness requirement on definitional or tactical grounds from those who make it a matter of moral and political principle, even though these distinctions are foreign to the works we shall be studying. Some thinkers will turn out to be using just the first and/or the second sorts of arguments distinguished here. I am not interested in these thinkers, for reasons I shall give in a moment. Most thinkers, however, not having made the distinction in question, will themselves be unclear as to just what they're defending, with respect to paying the penalty, much less why they are defending it. These are the thinkers I'm interested in. For after we have separated out the definitional and tactical components of their theories, we shall find in some cases that a willingness to pay the penalty remains an essential feature in the justification of civil disobedience according to these thinkers.

3. Perhaps Professor Hugo Bedau is the best example of the first sort of defense of the willingness requirement distinguished above. Bedau says at the outset of a justly famous article on the subject that he has been "unable to find a suitably detailed analysis of what civil disobedience is,"⁶ and that consequently he has decided to devote most of his energy to trying to provide an "analysis" of the concept of civil disobedience. It is in that part of his article devoted to his analysis or definition of civil disobedience that Bedau remarks that "normally, committing civil disobedience does not involve acting with...intent to resist, even non-violently, the legal consequences of the act"⁷ and calls the views of Irving Kristol and Robert Penn Warren "quite remarkable in implying the opposite".⁸

If by making acceptance of the legal consequences part of his

analysis or characterization of civil disobedience Bedau means to define a special sort of disobedience to the law, a kind of disobedience the very definition of which includes a willingness to pay the legal penalty, then obviously I have no quarrel with him. Professor Bedau, like everyone else, is perfectly free to define special categories of acts whose very definition includes this or that. Thus, we might think of Bedau as saying something like the following. There are many different ways of disobeying the law and many different reasons for doing so. One form of disobedience is of special interest to political theorists. This is a form of disobedience that is conscientious, nonviolent, and publicly performed, among other things.⁹ Let us call disobedience of this sort "civil disobedience". It's not that other sorts of disobedience cannot be justified -- they can. Thus, revolution is sometimes justified, and it may be quite violent. Alternatively, a clandestine, illegal abortion might be justified under certain circumstances. The point is not one about justification; it is simply one about what we are going to call the illegal act. And violent revolutionary acts, as well as clandestine abortions (where abortions are illegal), are not acts of civil disobedience as we are speaking of them, even when they're justified.

I think the preceding is pretty much what Bedau has in mind in his analysis of civil disobedience. The difficulty arises with the willingness requirement. It is not part of his explicit definition of civil disobedience, but I think Bedau more or less wants it to be.¹⁰ That is, he pretty clearly wants to say that whether or not one is ever justified in refusing to accept the legal consequences of illegal conduct, one is not performing "civil disobedience" if one so refuses. Thus, the willingness requirement becomes as much a part of the characterization or definition of civil disobedience as the requirement that it be public, say, or nonviolent.

In short, Bedau seems to be saying that when you break the law, even if you do so conscientiously, nonviolently, publicly, and so on, you are not performing "civil disobedience" unless you are also ready and willing to accept the legal consequences for what you're doing. This is all very fine, and Bedau is entitled to define civil disobedience in this

way. But there is one thing Bedau cannot do.¹¹ He cannot, without further argument, say that an act of disobedience to the law, exactly like an act that he would regard as an instance of justified civil disobedience and performed under exactly the same circumstances with exactly the same aims, is not justified if the disobedient is unwilling to suffer the full legal consequences of his action. To be sure, it is not an act of "civil" disobedience, on Bedau's view. And if the matter ended here, there would be little more to be said, unless one wished to quarrel about exactly how we ought to define or characterize civil disobedience, which I do not wish to do. But many thinkers do say, or imply, not only that an act exactly like an act of justified civil disobedience would not be civil disobedience in the absence of a willingness to pay the penalty, but also that such an act would not be justified. This view, it seems to me, is mistaken. Before proving this, however, I must dispose of another preliminary matter.

4. Many recent theorists (as well as practitioners) of civil disobedience have pointed out that if the disobedient is honestly seeking to effect substantive social change by his actions, then he should be prepared to suffer the consequences of those actions -- the legal consequences in particular. But this "should" is a prudential "should": the disobedient should (or ought to) express a willingness to accept the consequences because in many cases he is likelier to move others and achieve his aims by doing so.

This is certainly not an unimportant point. Civil disobedience is not moral posturing; it is, in the sense we are concerned with, an attempt to affect social policy, usually by affecting the majority's "sense of justice", as John Rawls has put it. And doing this will often involve important tactical considerations: How will the majority react to this particular act at this time in such and such a place? Will they take us seriously? Will they appreciate our sincerity and depth of feeling about this issue?

As many recent writers have pointed out, a willingness to accept the legal consequences of one's action is particularly helpful with

respect to the last question above. Thus, Professor Marshall Cohen has written that:

After openly breaking the law, the traditional disobedient willingly pays the penalty. This is one of the characteristics that serves to distinguish him from the typical criminal... and it helps to establish the seriousness of his views and the depth of his commitment as well.¹²

Cohen is quite emphatic in denying that accepting the penalty has anything to do with the justification of the illegal act and in emphasizing what is (for him) its strictly tactical relevance. Thus, speaking of the suffering that accepting the penalty often entails, he says:

"This suffering does not justify the act of civil disobedience, but it helps to establish the disobedient's seriousness and his fidelity to law in the eyes of the majority...".¹³ In the same vein, Cohen suggests still another tactical reason for being willing to suffer the legal consequences of an illegal act:

The disobedient's willingness to suffer punishment has another purpose as well. It is meant to weaken the will of the transgressors and to discourage them from a course of action that the dissenters consider immoral. For, if the transgressors do not draw back, they may be forced to punish some of the most scrupulous and dedicated members of the community. The fact that this is so will often persuade those who heedlessly supported the original measures, not to mention those who supported them with a dim sense of their injustice, to withdraw their support or even to join the opposition. Forcing others to suffer for their moral beliefs is a high price for pursuing a questionable course of conduct and many will prefer not to pay it.¹⁴

In what follows it is not part of my intention to take issue with these and similar tactical considerations. Indeed, as I hope I have made clear, I agree with Professor Cohen that a willingness to accept the legal consequences of one's (illegal) acts may often be a very important tactical consideration in perpetrating effective civil disobedience. Hence, if an author emphasizes the importance of accepting the legal consequences of

an illegal act because of the strategic or tactical effect such an acceptance might have -- in convincing the general public of the sincerity of the disobedients, for example, or of the purity of their motives -- then I have no quarrel with him, for he is not making a moral issue out of the willingness requirement in the sense that interests me.

But we must not confuse tactics with principles (Professor Cohen certainly does not, as we have seen), and the position I shall be concerned with is one that insists on a willingness to accept punishment (or any other legal consequences) not for tactical reasons but for reasons of principle -- namely, because the illegal act is not morally justified unless the agent is willing to accept its legal consequences. I think it is fair to say that the latter view is the position of the thinkers whose views we shall now consider, and where it is not, I am not concerned with attacking them.¹⁵

5. The clearest kind of case in which we would have an example of a thinker for whom the willingness requirement is a matter of moral and political principle in the justification of civil disobedience is that of a thinker who explicitly claims that a willingness to pay the penalty is one of the necessary conditions for the justification of an act of civil disobedience. Professor Stuart M. Brown provides us with just such an example. He writes:

In order to be justified, acts of civil disobedience must meet each of at least three conditions: (1) persons may not be harmed, and property may not be destroyed; (2) there must be unconditional submission to arrest and to the legal penalties for the breaches; and (3) the protests, in the course of which the breaches occur, must be directed at constitutional defects exposing either all the people or some class of the people to legally avoidable forms of harm and exploitation.¹⁶

It is, of course, the second of these three conditions for the justification of civil disobedience that interests us here. Unfortunately, Brown does not provide an independent defense of this second

condition; instead, for reasons he does not explain, he first defends his third condition at great length and then returns to the first two conditions and argues for them together. After explaining and defending the third condition, then, Brown writes:

The remaining two conditions for justifying civil disobedience are the protection of persons and property and the passive submission to arrest and punishment. The first restricts the kind of law that may be broken; the second restricts the extent to which the defiance of a given law may be carried.¹⁷

Brown then gives three arguments for these two conditions, without explicitly indicating to what extent each argument specifically applies to one, the other, or both of these conditions. He begins with the following argument:

There are three reasons for making these restrictions. The first is that the use of force and violence, being evil in itself and being no less evil for being used in a good cause, can be morally justified only in circumstances where the alternative is an even greater evil and cannot, therefore, be justified in cases of civil disobedience.¹⁸

As it stands this argument involves a patent non sequitur: it simply does not follow from the fact that force and violence are great evils, which can only be justifiably used in special circumstances, that force and violence cannot be justified in cases of civil disobedience. Indeed, some recent writers have been inclined to think that property is less than universally sacred, and while they have been anxious to retain the prohibition against violence toward persons, violence toward property is sometimes thought to be justifiable in at least some cases of civil disobedience.¹⁹ We need not go into this here, however, for although Brown has led us to believe that his argument is directed at both the first and the second conditions, this first reason for the restrictions those conditions imply bears only on condition (1), which protects "persons and property" and forbids "the use of force and violence."²⁰

Brown comes to the point when he gives his second and third reasons for conditions (1) and (2). He says:

The second reason for imposing these restrictions is the need to maintain a clear, sharp distinction between justified acts of civil disobedience and justified acts of civil rebellion.²¹

This argument is certainly obscure -- at best enthymematic -- but I think Brown's point is this: violence and an unwillingness to accept the legal consequences of one's (illegal) acts may be justified in all-out rebellion, but we must be careful to distinguish such civil rebellion from "civil disobedience", which is a much milder form of protest and where violence and this unwillingness cannot be permitted.

Now if this were merely what I have called a "definitional" move on Brown's part, I would have no quarrel with it. However, as I shall show shortly, there is conclusive evidence (indeed, an explicit statement) in Brown's paper that his reasons for (1) and (2) are not intended to be merely definitional. But if this is so, then this second reason simply rests on a confusion. For the idea is that, while violence and an unwillingness to accept punishment may be justified in an all-out rebellion, we must still preserve an alternative form of civil protest that is nonviolent and where the dissident is willing to accept the legal consequences of his acts. But this is to assume that there is only one alternative to all-out civil rebellion, which is obviously false. The possibilities for protest and dissent range from something like a mild voicing of polite disagreement with the law or policy in question, through "traditional" civil disobedience, so-called "direct action" and so forth, all the way up to full-scale revolution.

Now at any stage short of all-out revolution, for both tactical and theoretical (I mean "principled") reasons, violence and a willingness to accept the legal consequences of one's actions may or may not be justified. And even if at some point violence may not be justified (until conditions get worse), it does not follow that an unwillingness to pay the penalty is justifiable only when violence is justifiable. As we

shall see in Chapter II, such a claim assumes that one can display such an unwillingness only in some manner that involves force or violence, an assumption which is plainly false. In any case, Brown certainly cannot justify conditions (1) and (2) simply by calling for a clear and sharp distinction between "justified acts of civil disobedience and justified acts of civil rebellion." This is indeed an important distinction to observe and to preserve; it does not ensure conditions (1) and (2), as Brown would like it to (and needs it to), however, unless these are the only two courses open to a dissident.²²

Finally, Brown gives his third reason for conditions (1) and (2) and the "restrictions" they imply:

The third reason for maintaining these restrictions is to preserve civil disobedience as a tolerable, ritualized form of public protest in which law-breaking is minimal and for the most part formal.²³

Now when Brown argues for conditions (1) and (2) by saying that they will "preserve civil disobedience as a...ritualized form of public protest", and so forth, it almost seems inevitable that we conclude that he is now simply defining what he considers an important form of political protest that is well worth preserving. But such an inference flies in the face of Brown's explicit assertion that "a public protest in which the participants injure others, destroy property, and resist arrest is no less an act of civil disobedience; it is merely an unjustified one."²⁴ I think this assertion makes it clear that Brown does not intend his remarks in what I have been calling the "definitional" sense. But if this is not his intent, then what is it, since this third "reason" is certainly a difficult one to interpret?

I think the answer is that this last (third) reason is what I called earlier a "tactical" one -- but with a difference. Some authors insist on non-violence and a willingness to accept legal consequences as an important tactical consideration in certain particular instances of civil disobedience; but I think Brown's argument is slightly more subtle, or at least somewhat different. For he is arguing that it is

tactically important not just for any particular act of disobedience to be characterized by conditions like (1) and (2) but that, from the point of view of the disobedients, it is important for there to be an institution or institutionalized form of disobedience to the law characterized by such conditions.

The argument, as I understand it, rests on the claim that it is strategically important that there be some form of dissent which the general public can recognize as a form that will be both nonviolent and submissive (with respect to legal consequences), despite its illegality. The reasons for this latter claim are somewhat obscure. Is the point simply that if there is such a ritualized form of dissent, with non-violence and the willingness requirement built into it, then the public will be more likely to appreciate the sincerity of the dissenters and to distinguish them from run-of-the-mill criminals? If so, the argument faces a number of difficulties. First, it seems to take a rather cynical view of the general intelligence in a democratic society. I return to this point in Chapter II. Secondly, it is a tactical point, even if a generalized one, and hence does not prove -- what Brown had set out to prove -- that in order to be justified civil disobedience must, in principle, be characterized by a willingness to accept the consequences, much less that such a willingness is a necessary condition of justified civil disobedience.

But suppose we concede Brown's cynicism about the public's ability to assess different forms of illegal conduct differently, and suppose we further concede that, in light of this, "it is important," as he says, to preserve a form of disobedience to the law like that characterized by Brown. There is still a third difficulty in Brown's argument. For it now seems that Brown is making both a definitional and a tactical point: we ought to include nonviolence and the willingness requirement in our very conception of justified civil disobedience because it seems strategically important (from the point of view of the disobedients themselves) to do so.

It is difficult to evaluate such a claim in the absence of more

explicit arguments, and I shall not be able to do so in any detail until Chapter II, where I consider a related claim made more recently, and defended more cogently, by both John Rawls and Marshall Cohen. But it is important to notice here that by failing to separate the nonviolence requirement and the willingness requirement, Brown has left himself in an awkward position. There are many reasons for cautioning dissenters against violence -- indeed, Brown has mentioned some above -- and tactical considerations are not even foremost among them.

But when Brown insists that in order to be justified civil disobedients must express a willingness to pay the penalty, he must give us more than definitional or tactical rationalia for this claim (if he wishes to make more than definitional or tactical points, at any rate, as he obviously does).²⁵ For now he is no longer advising dissenters to respect the persons and property of others; he is insisting that they welcome damage (in the form of punishment and fines) to themselves. This they may be willing to do, for the sake of their long-range goals and if the circumstances seem to require it. But in the absence of arguments to the contrary, it would seem that this decision, being a tactical one, should be left to the dissenters.

What such "arguments to the contrary" would have to show, of course, is why, as a matter of moral principle and not simply as a matter of political cunning, a dissenter has to pay the legal penalty for resisting unjust legislation. Brown does not provide us with such arguments. But this third argument is a suggestive one, and I return to it at the beginning of Chapter II when I consider the view of John Rawls and Marshall Cohen.

6. Brown's paper appeared when the decade of discussion on civil disobedience was just beginning, and it would be unfair to expect him to see and effectively deal with problems that arose precisely because of the work of thinkers like himself. Before we proceed to the arguments of more recent thinkers, however, it will be instructive to look at the position of one other early theorist of civil disobedience, Professor Milton Konvitz of the Cornell University Law School. The principal merit of

Professor Brown's position is that he says quite straightforwardly that the willingness requirement is a necessary condition of justified civil disobedience on his view. Unfortunately, Brown's arguments for this claim are rather sketchy, as we have seen. Professor Konvitz, on the other hand, provides arguments of some length and subtlety in defense of the willingness requirement. But Konvitz is like many more recent thinkers in an unfortunate respect: he does not say explicitly that the willingness requirement is a necessary condition in the justification of civil disobedience.

There is no question that Konvitz is defending this requirement for reasons other than the tactical and definitional ones distinguished above, as we shall see. And he defends the moral basis of the willingness requirement at some length. But what we gain in depth of argument, we lose in clarity with respect to just what role the willingness requirement is playing in Konvitz's theory of civil disobedience. This is a problem we shall have to face again and again in looking for arguments for and defenses of the willingness requirement. It is not until we get to the theory of Professor John Rawls in the next chapter that I shall be able to clear up these difficulties completely.

Konvitz has written that "voluntary -- and even willing -- assumption of the legal punishment for violation of the law is an essential characteristic of civil disobedience."²⁶ This could easily amount to no more than the claim that disobedience to the law is not "civil disobedience" unless the disobedient willingly accepts the legal consequences of his action. But in the context of Konvitz's article it becomes clear that he wants to say more than this, although it is less clear just how much more he wants to say. Thus, he claims that civil disobedience "entails the penalty of the law" and that "the penalty must be imposed", adding cryptically that "in this way [i.e., by straightforward imposition and willing acceptance of the penalty], civil disobedience may render unto Caesar what is Caesar's and unto God what is God's".²⁷ As if to underline his position with a final thrust, Konvitz concludes his article by remarking that:

by looking only at the act of disobedience, one sees only half of the event. The punishment is the other half, which is just as essential to the person as the act of disobedience.²⁸

Why is all this so, according to Konvitz?

I think we can distinguish three different reasons for this view in Konvitz's paper. These lines of argument are at least hinted at in a great deal of other writing on this subject, and it will be useful to deal with them once and for all here. Only the first of these three lines of thought has much plausibility, but the other two are worth sketching and refuting. Konvitz does not himself distinguish these three different reasons for his view, and in order to sort them out I shall have to quote three different passages from his article.

It will be easier to see the point of Konvitz's first argument for the willingness requirement if we bear in mind that in his paper Konvitz is replying to Rawls's early paper "Legal Obligation and the Duty of Fair Play".²⁹ Rawls had not yet articulated his subsequent theory of the justification of civil disobedience, much less his defense of the willingness requirement, but Konvitz was quick to anticipate at least part of it. In the passage below Konvitz speaks of the intention, as he sees it, of the civil disobedient and describes it in terms of Rawls's concept of "fair play".

- I. The intention of his [the disobedient's] act is not to override the principles of fair play but to achieve their vindication in the wills and institutions of men; and I believe that fair play requires that we do justice to this intention in our attempt to formulate the moral principles that are the underpinnings of obedience and disobedience. Such men cannot be in any way charged with trying to be "free riders". They try, through civil disobedience and paradoxically, to affirm, rather than to breach, the agreement between themselves and the Laws of which Socrates spoke.³⁰

Now in the suggestion (above) that the conscientious law-breaker is trying to affirm rather than to breach the "agreement" between himself and the

Laws in disobeying the law in question, I think we have the seed of the only of Konvitz's three reasons for his view that is at all possible. And it is significant, I think, that this particular reason is couched in "contractual" terms -- i.e., in terms of the hypothetical (or fictitious) "agreement" between the law-breaker and the Laws. However, Konvitz does not develop the line of argument that is implicit in this passage until the last paragraph of his article, where he only adumbrates it quite sketchily, and instead proceeds to an entirely different and rather curious line of thought and then to an even more curious one. The first of these two is contained in the following passage, a small part of which I have already quoted:

II. What Thoreau said, and what Professor Rawls is saying, is that the moral judgment must be given priority over the political or any other judgment. When pressed far enough, this means that the individual conscience must have the last word. If that last word means civil disobedience, which entails the penalty, then the penalty must be imposed, and in this way both conscience and law are vindicated. Socrates could not and would not accept as binding on his conscience the majority vote of the jury that tried him, but he willingly submitted to the penalty. So, too, a citizen may not accept as binding on his conscience the majority vote of the legislature, though he will willingly submit to the penalty for his civil disobedience of that law. In this way civil disobedience may render unto Caesar what is Caesar's and unto God what is God's.³¹

Now it simply seems to me absurd to suppose that conscience is in any sense "vindicated" simply by allowing disobedience that is summarily punished. The "Caesar-God" metaphor, moreover, only serves to make matters worse. Konvitz says that "in this way," i.e., by both calling the disobedience "justified" (and in that sense "allowing" it and supposedly giving conscience "its due") and then promptly punishing it, we can see to it that civil disobedience renders "unto Caesar what is Caesar's and unto God what is God's". Now Caesar may well be getting his due in such an exchange, to continue Konvitz's figure, but it hardly seems that God (or conscience) is. For how could He? A man has acted,

ex hypothesi, as God wants him to act. Quite clearly, neither God nor the man can be expected to feel that he's "gotten his due" when the latter is subsequently punished for doing exactly what the former wanted him to do. At any rate, I doubt that God, if consulted about the matter, would take Konvitz's sanguine view about whether He's getting His due.

In any case, however, I think that this whole way of conceiving the matter (as expressed in II) is a muddle that Konvitz has fallen into, and that at least in Passage I, where he speaks about the intention he thinks the disobedient is acting out of, Konvitz is on to a somewhat more plausible line of argumentation. He returns to this line of thought in Passage III below, but not, unfortunately, without mixing in another and even less convincing "reason" for the willingness requirement.

III. By looking exclusively at the act of disobedience, one sees only half of the event. The punishment is the other half, which is just as essential to the person as the act of disobedience; for even as he breaks the law, he intends to restore its wholeness. He may be mistaken in what he does, but he ought not to be blamed for wrongs that are entirely beyond his intention. "What will thou have? quoth God; pay for it and take it." This Emersonian rule is the measure by which he who breaks the law in civil disobedience acts. He takes and pays, and thus affirms, rather than overrides, the duty of fair play.³²

If we remember that the key-note of Passage I was the disobedient's affirmation of the agreement between himself and the Laws, we can see that Konvitz is returning to this line of thought in Passage III. At best, however, III contains two kinds of reasons for the willingness requirement -- one very bad one and another not so bad (though ultimately unsuccessful) one. At worst, it contains only one reason: the "not so bad one" supported by (cf. the 'thus' in the last sentence of III) the really bad one. Let me explain.

On one interpretation Passage III contains two different lines of thought in support of the willingness requirement. The first is a development of that which was adumbrated in Passage I: in breaking

what he considers an unjust (or immoral) law, and accepting the legal consequences, not only is he affirming the agreement between himself and the Laws, but the conscientious law-breaker is also seeking to restore the "wholeness" of the law in breaking it. I shall have more to say about this line of thought below.

The second line of thought in III, on this interpretation, and Konvitz's third reason in support of his view, is embodied in the "take-and-pay" metaphor. The idea, presumably, is that paying the penalty is part of the cost of breaking the law; and the disobedience is justified only if the disobedient is willing to pay that penalty. Professor Marshall Cohen has expressed the absolute unacceptability of this view far better than I can hope to do:

It is in interpreting the role of punishment in the theory of civil disobedience that many errors are made. For the theory of civil disobedience does not suggest...that the disobedient's actions are justified by his willingness to pay the penalty that the law prescribes.... This suffering does not justify the act of civil disobedience, but it helps to establish the disobedient's seriousness and his fidelity to law in the eyes of the majority whose actions have, in his opinion, justified it.³²

Now what are we to make of Konvitz's other view that "punishment is just as essential to the person as the act of disobedience" itself, because "even as he breaks the law, he intends to restore its wholeness", or because such men are trying "through civil disobedience and paradoxically, to affirm rather than to breach, the agreement between themselves and the Laws of which Socrates spoke?" Certainly, it is admirable that it should be part of a conscientious law-breaker's intention to affirm his agreement with the Laws and to try to restore the "wholeness" of a law that is, to him, for moral reasons or for reasons of justice, somehow "out of joint". But while agreeing that such intentions are indeed admirable, we must be careful to ask (i) how it is that his having these intentions requires him to invite (or at any rate willingly accept) punishment, or any other legal consequences of his act; and (ii) whether, indeed, such intentions are in any sense necessar-

ily connected with justified disobedience to the law.

I think it is clear from Konvitz's article that the answer to (i) is a tactical one and not a matter of any moral principle attaching to the theoretical justification of the act in question; that is, a willingness to accept the legal consequences of his act may be good tactics just because it may help the disobedient to distinguish himself from the run-of-the-mill criminal, to establish his sincerity in the eyes of the rest of the community, indeed, perhaps help "win them over", and so forth. As I hope I have made sufficiently clear earlier, however, such tactical considerations in favor of a willingness to accept punishment and other legal consequences in no way constitutes an argument for making such a willingness a necessary condition of justified disobedience to the law.

Hence, Konvitz's claims that "voluntary -- and even willing -- assumption of the legal punishment for violation of the law is an essential characteristic of civil disobedience" and that "...civil disobedience...entails the penalty of the law" simply stand unproven -- unless, of course, he is simply stipulating that this is a necessary feature of a certain species of disobedience to the law called 'civil disobedience'. It is abundantly clear from Professor Konvitz's article, however, that he is not merely making such a definitional move. But since (or if) he is not, then he needs arguments for the willingness requirement and he simply does not offer them; or rather, he offers two unsuccessful arguments for his position and one sound argument to the effect that conscientious law-breakers should be sensitive to tactical considerations.

One final word about this third argument of Konvitz's, that a willingness to accept punishment is required as a sign of the disobedient's affirmation of the "agreement" between himself and the Laws or as an attempt at restoring the "wholeness" or integrity of a bad law. It may be that in certain cases of the kind we are considering we will need "signs" or even explicit assurances that the disobedient is not trying to rupture the fabric of the law entirely or to break the "agreement" in question. But it seems slightly paradoxical to insist so

stringently on signs and assurances of this sort when we remember the kind of case we're dealing with -- namely, cases in which it is precisely the disobedient's claim that this "agreement" has been broken by others (society, say, or some part of it, or some one of its laws or policies). This man certainly doesn't need to be reminded of the "agreement" or the principles of "fair play", since it is exactly his claim that the former has been broken or that the latter have been violated.

It is in the context of the preceding remarks, I think, that we can answer question (ii) above quite simply. There is, of course, a "necessary" connection of some sort between intentions (on the part of the law-breaker) of the kind Konvitz describes and justified disobedience of the kind we are considering. For it is precisely the spirit of civil disobedience as we are construing it that the "agreement" between the individual citizen and the Laws (as Konvitz puts it, following Socrates in the Crito) must not be broken. It follows analytically from the fact that a man chooses to defend disobedience to the law with an argument like Socrates' "contractarian" argument in the Crito that he has intentions in some sense like those Konvitz wants him to have. Hence, it is all the more paradoxical to suppose that we need to "make sure" he has such intentions by insisting that he accept the legal consequences of his action. Indeed, if his case can be made good, along the lines suggested by Socrates and elaborated recently by John Rawls (see Chapters III and IV below), then it would in fact be wrong (from the point of view of "morality", "justice", "utility", and so forth) to punish him, much less to expect him to invite punishment. Before I can show this, however, and thus give Konvitz's "fair play" argument its due, I shall have to sketch Rawls's theory of civil disobedience as well as his arguments for the willingness requirement. Before doing that there is an entirely different kind of defense for the willingness requirement that I want to consider.

7. According to the willingness requirement it is morally right for society to punish an individual who is, ex hypothesi, doing the morally right thing in disobeying the law, and it is morally incumbent upon that individual to accept that punishment and to do so willingly. This is

paradoxical, but I must insist that it is part of the received view on the justification of civil disobedience. After all, a theory of civil disobedience is supposed to tell us when civil disobedience is morally justified. And according to many recent thinkers civil disobedience can be justified, even in more or less "just" democratic societies, but only if the disobedient is willing to pay the penalty for his disobedience. It is this conjunction of (i) attempting to show that under certain circumstances it can, on balance, be morally right (and perhaps obligatory) to perpetrate civil disobedience, while (ii) insisting that it is nonetheless morally right (and perhaps obligatory) for society to punish such disobedience, that generates the paradox that worries me.

Perhaps it is not inappropriate to point out that resolving this (apparent) paradox involves us in much more than an academic dispute. For if what I call the received view is correct, then civil disobedients who, perhaps understandably, do not want to pay the penalty cannot justly accuse society of compounding injustice with injustice when they are punished for violating an admittedly unjust law. But if I am correct, and the received view is unsound, then in inflicting punishment on individuals who have justifiably broken the law society is treating them unjustly and is itself perpetrating a serious moral wrong. And to the extent that the received view rationalizes this injustice it is not merely guilty of intellectual error but of encouraging moral error and injustice as well.

So much might be said for the importance of our inquiry. But these remarks have another bearing on our problem as well, for they suggest what one would have thought would be the most natural and straightforward way of determining whether it is morally right to punish people who are themselves morally right in what they've done. Curiously, until quite recently, no one I know of had taken this approach. In fact, I had attempted (and failed) to work out a possible defense for the willingness requirement on my own, in the way considered below, when yet another book on civil disobedience was published which, if it has no other merit, has the distinction of being one of the few works in the recent literature that takes the question of justifying the willingness requirement (and the

punishment that comes with it) at all seriously. What's more, it does this in a rather natural way: by asking whether the principles that allegedly justify inflicting punishment for ordinary criminal acts also justify punishing those whom we agree are not only conscientious in their disobedience but morally right in perpetrating it as well.

The traditional defense of the institution of punishment, as well as of its application to particular cases, proceeds along well-known lines. There is certainly not much agreement as to just what justifies us in administering punishment to law-breakers, if we're justified at all, but there is fairly widespread agreement on the possibilities. Most attempts to defend punishment rely on the rationale of "deterrence", as it is called, some rely on that of reform or rehabilitation, a few on the notion of "retribution" or some more sophisticated variation on retribution; and of course, many rely on some combination of two or even all three of these traditional rationalia. It seems natural to ask, then, whether any of the standard arguments for the general institution of punishment, either by itself or in conjunction with one or both of the other arguments, can be used to defend the willingness requirement. And even if we are skeptical about these traditional moves in defense of punishment generally, we might ask whether the case for punishing those who have justifiably disobeyed the law is any worse than the case for punishing law-breakers generally. Professor Carl Cohen has recently tried to show that it is not and that the rationalia of retribution, reform and deterrence apply just as much to justified civil disobedience as to unjustified, run-of-the-mill law-breaking.³³

Perhaps the most widely-received argument for saying that punishment is at least sometimes morally tolerable is based on the theory of deterrence. On this view punishment is seen as a way of discouraging the disobedient from further illegal action as well as discouraging others from following the disobedient's example. I wish to disregard for now the critically important claim that in fact punishment is far less effective as a deterrent than is commonly thought. Suppose we grant that for certain crimes and certain types of criminals (and potential criminals) deterrence is to some degree effective and hence not an altogether unlikely

candidate for justifying punishment in at least a wide range of cases. Nonetheless, even if we accept deterrence as a reasonable rationale for punishment in some general sense, it is not obvious how the argument is to proceed when we are talking about punishing people for crimes that we are willing to concede were themselves morally right. The obvious question one is tempted to raise is simply this: why should we want to deter people from acting in a way that we agree is morally right despite the fact that it is clearly illegal? If we assume, as Carl Cohen himself does, that an act of civil disobedience is morally justified when it produces more good for society, in the long run, than evil, it seems odd to say that such conduct should be discouraged.³⁴

But of course an obvious reply is available to Cohen, and he attempts to make use of it. Ideally, it might be true that we should not punish those who have done more good for society by disobeying the law than they would have done by obeying it, or when they have disobeyed the law in order to avoid the injustice of obeying it. Presumably we actually want to encourage such conduct. But if the individual who (justifiably) engaged in civil disobedience is not punished, will this not encourage others, who not so correctly judge the long-range "good" or the immediate "justice" of a statute, to act in a similar manner? And surely in allowing this we are courting chaos and disaster. For the example set by the disobedient may be misconstrued, and when he is not punished, many people who would otherwise be deterred from unjustifiable disobedience may feel justified in acting illegally and may actually do so, when in fact their disobedience is not justifiable -- i.e., does not in fact contribute more good than evil to society in the long run, or is not necessary in order to avoid perpetrating injustice by virtue of complying with an unjust law.³⁵

This is a curious argument. For one thing, it rests upon an important empirical claim: I call it the "domino theory of civil disobedience." In this regard one would simply like to know the facts. For example, is it true, as a matter of empirical fact, that ordinary people are unlikely to perceive the difference between conscientious resistance to segregation laws, say, and other sorts of criminal conduct? I know of no recent research that would shed any light on questions like this one, and Cohen

does not allude to any. This is significant, since his argument depends on the truth of something like the "domino theory" of civil disobedience.

But this argument faces an even more serious difficulty. For even if it could be shown that exempting justified civil disobedience from the ordinary routine of punishment would have a certain amount of "negative utility", to put it crudely, surely we would want to balance this against the apparent injustice of punishing the man who has done the right thing, from either a "teleological" or a "deontological" point of view, or both, in disobeying the law. Such a person's claims to consideration in the calculus of good and evil is surely not insignificant. Moreover, on this view of things, we should have to take seriously the possibility of a reverse "domino" effect. That is, we should have to ask how much socially beneficial disobedience we would be discouraging by punishing offenders whom we are willing to concede were right in what they did. A society that punishes such people might lose more in the long run than it gains, not simply by discouraging further socially useful disobedience but also by slowly undermining its own moral standing in the eyes of many of its members.³⁶

The most that can be said, then, for the "deterrent" argument in favor of punishing even disobedience that is admittedly justified, is that it rests on a dubious empirical claim and an even more dubious assessment of moral priorities. The whole notion of "justified" civil disobedience rests, among other things, upon the fact that some acts of disobedience are better for society than not, and that some are a consequence of our moral commitment to "justice" and "right conduct" generally. If it is conceded that in fact we do not want to deter such conduct, but only the unfortunate consequences of ignorant misconceptions of what justified civil disobedience is all about, then deterrence theory will not provide much succour to proponents of the willingness requirement until they come up with the appropriate empirical data and, even more importantly, until they show us that the evil of punishing a man who has acted rightly in breaking the law is outweighed by the evil allegedly engendered by the example he sets in acting as he does. At present, it seems extraordinary to me that thinkers like Carl Cohen are so confident that civil disobedience sets a

"bad" example, even when it is justified, and that they are willing to defend punishment of the righteous man before any of the facts are in and before the calculus of social good has shown that he must, unfortunately, be punished for the common good. It is perhaps an understatement, then, to say that the burden of proof is on Cohen (and his allies) here, and that this burden has not been sufficiently discharged as yet to warrant Cohen's claim that deterrence theory justifies punishment of justified civil disobedience just as well as it justifies punishment in general. This is not to say that Cohen's unsubstantiated empirical claim is absurd or altogether implausible on its face. It simply stands unproven as things are.

Another traditional rationale for the institution of punishment is rehabilitation or reform. On this view society "punishes" the disobedient in order to reform him, or change his perspective on anti-social behavior, so that his conception of social behavior will be more in accord with that of the society at large. At least up to the point, at any rate, where he will no longer be inclined to engage in outright anti-social behavior. Here there is even less reason to believe that our present penal system is achieving this goal, with any degree of effectiveness, than there is to believe that it is achieving the goals of deterrence. But suppose we avoid this difficulty for now and pretend that we can in fact "reform" criminals and make them law-abiding citizens. We still face a question similar to the one raised above: why should we want to "reform" an individual whom we agree acted rightly in disobeying the law? What rehabilitation is in order for the man or woman who rightly protested racial inequality, for example, or any other form of legalized injustice?

Carl Cohen does not have even the hint of an answer to this question, which is not surprising. There was at least a superficially attractive way of explaining why we might want to deter people from attempting to perpetrate justified civil disobedience. But no similar move is available here, for the argument would have to contend that for some reason we actually want (and need) to reform people whom we also want to encourage and applaud. It is hard even to imagine what such an argument would be like, unless it took the form of saying that we wanted (and needed) to eliminate civil dis-

obedience from society altogether. But neither Cohen nor any other proponent of the received view can say this, since it is precisely their contention that civil disobedience can, in theory, be morally justified, and is, in fact, a form of dissent that we both need and want in our society. Hence, the rationale of reform or rehabilitation lends no support to the willingness requirement whatever, even if we grant it a fragile plausibility as a rationale for the punishment of ordinary criminal behavior and even conscientious but unjustified civil disobedience.

There is, finally, the so-called "retributive" rationale for the institution of punishment. It is perhaps oddest of all to suppose that this objective would justify punishing justified civil disobedience, even if it can justify avenging cold-blooded murder, say, with cold-blooded capital punishment. Again, the question one would want to press here is obvious: just what are we exacting retribution for when we punish a law-breaker who was morally justified in what he did?

Surely not for "doing wrong"; for, ex hypothesi, he "did right". And it is certainly not open to the retributivist to say that, despite the moral rightness of what he did, the civil disobedient must be punished because, after all, he broke the law, although Carl Cohen tenders this reply and makes much of it.³⁷ Such an argument contradicts what is perhaps the single component of retributivist theory that carries even the slightest plausibility: that the moral propriety of punishment is somehow related to the moral impropriety of what was done and that somehow it is appropriate, morally speaking, to do evil to evil-doers (to extract an eye for an eye and a tooth for a tooth). It is not the bare illegality of an action that moves the retributivist to his wrath but the immorality or downright "evil" of that action. Since the case we are concerned with is that of an individual who acted rightly in what he did, it is impossible to tender an argument whereby punishment would be justified on the grounds that somehow we were "righting" a moral wrong. In the cases we are dealing with there is simple no moral wrong to set right, in the retributivist sense.³⁸

This is as much as I want to say for now about the sort of approach

Carl Cohen takes with respect to the justification of the willingness requirement and the punishment of justified civil disobedience. I do not pretend that I have said the last word on these matters, and I shall return to them in Chapter V below. I do think, however, that enough has been said to show how difficult it would be to defend the willingness requirement on the same grounds that we defend punishment in general. Carl Cohen has certainly not shown how this can be done, and in Chapter V I shall try to show that this is not merely due to his particular limitations. It is due to the fact that any attempt to vindicate the willingness requirement, while granting that civil disobedience can be morally justified, is in principle misconceived; for the received view is incoherent in as much as it makes the willingness requirement a part of the justification of civil disobedience.

CHAPTER II

RAWLS'S ARGUMENTS FOR THE WILLINGNESS REQUIREMENT

1. I wish to turn now to some very recent work on the role that paying the penalty should play in the theory of civil disobedience. It will be helpful to begin by recalling a difficulty we got into with respect to Professor Brown's views in this regard. This will involve a certain amount of "backtracking", but it will enable me to connect this chapter with the previous one and, more importantly, will put the present discussion in the proper perspective.

In one of his arguments for the restrictions that make nonviolence and a willingness to pay the penalty part of the justification -- and not simply part of the definition or tactics -- of civil disobedience, Brown wrote:

The third reason for maintaining these restrictions is to preserve civil disobedience as a tolerable, ritualized form of public protest in which law-breaking is minimal and for the most part formal.

I think Brown's third argument can be instructively compared with some of Professor Marshall Cohen's arguments in a recent article on "Liberalism and Disobedience". Getting clear on Cohen's view, as compared to Brown's, will enable me to identify the exact position with respect to paying the penalty to which I am opposed and against which I shall argue. I'm fairly certain that Cohen himself does not hold the view I wish to rebut. However, I think I can show that John Rawls does. Unfortunately, it is not perfectly clear that even Rawls holds exactly the view I am opposed to, but it seems to me that he holds some view close to it. In any case, I am only interested in criticizing Rawls's (or anyone else's) position with respect to paying the penalty to the extent that he makes a willingness to accept the consequences of civil disobedience a matter of

moral and political principle in the justification of such conduct.

It will be recalled that in trying to get clear on the arguments involved in Brown's third reason (above) for the nonviolence and willingness requirements, I noted that Brown was in danger of defending these requirements on what I call merely definitional or tactical grounds and not on grounds of moral and political principle. That is, he seemed almost to be saying that it was important to have a ritualized form of dissent whose very definition included the notions of nonviolence and a willingness to pay the penalty, such that acts which did not have these characteristics, even if justifiable, would not be "civil disobedience". What's more, he seemed to favor this definition (or characterization) of civil disobedience for some sort of general tactical reasons (see section 5, Chapter I). But this interpretation seemed unfortunate for Brown's argument because he says explicitly that "a public protest in which the participants injure others, destroy property, and resist arrest is no less an act of civil disobedience; it is merely an unjustified one." There seems to be some inconsistency here. On the other hand, it might be argued that my division of the arguments for the willingness requirement into "definitional", "tactical", and "moral-and-political-principle" arguments is an artificial one and, what's worse, is a division that obscures the merits of some of the best arguments on behalf of the willingness requirement as well. I want to explore this possibility before addressing myself to Rawls's defense of the willingness requirement.

I noted earlier that Professor Marshall Cohen is extremely sensitive to the tactical relevance of paying the penalty. Thus, in a recent article he has written:

After openly breaking the law, the traditional disobedient willingly pays the penalty. This is one of the characteristics that serve to distinguish him from the typical criminal (his appeal to conscience is another), and it helps to establish the seriousness of his views and the depth of his commitment as well.¹

Cohen expands upon these remarks somewhat and finally concludes: "The

disobedient's willingness to face suffering and punishment may be seen, then, as a useful way of reinforcing the effects of his protest and appeal."² It would be a mistake to conclude from these remarks, however, that Cohen's sole (or even principal) reason for urging disobedients to pay the penalty is a tactical one. He considers at least two other reasons as well. One of these, which we shall consider below when we discuss Rawls's views on the willingness requirement, has to do with the notion of "fidelity to law" and its place in a theory of civil disobedience designed to fit a reasonably just constitutional democracy.³ As I shall show, there is some reason to believe that Rawls himself believes that under certain conditions not only is illegal dissent without a willingness to pay the penalty not "civil disobedience" but that, under those conditions, it is not justified either. But Cohen does not seem to want to say this. In remarks addressed to a criticism made by Professor Kai Nielsen, Cohen says that "Nielsen is essentially correct ...in thinking that I am putting forward a definite conception of civil disobedience,"⁴ and later continues:

In defining civil disobedience in such a way that the civil disobedient necessarily accepts the punishment, I have in no way committed myself on the question of [moral] legitimacy. It certainly remains open to me to say that there are justifiable acts of illegal dissent that do not require the acceptance of punishment. The definition offered simply has the consequence that where the disobedient does not accept punishment (and conduct himself in a number of other specified ways), his actions will not count as civil disobedience in the sense sketched here, and they will not be subject to the observations made about that form of politics.⁵

This is about as clear a case of what I call the "definitional" defense of the willingness requirement as I can imagine. Thus, Cohen illustrates very nicely the position of both the thinker who urges a willingness to pay the penalty for tactical reasons and the thinker who defends the willingness requirement as part of the very definition of civil disobedience. What's important about the latter point is that defining civil disobedience in one way rather than another is not

a matter of mere caprice, as Cohen makes quite clear. Rather, there is a certain form of illegal activity that not only has a tradition of its own in recent democratic politics but that is also particularly appropriate as a form of dissent in more or less "just" democratic societies.

When I dismiss a defense of the willingness requirement as "merely" definitional or tactical, then, I do not mean to overlook both the tactical reasons for accepting punishment in a particular case and, perhaps more importantly, the more general tactical reasons for having a generally recognized form of disobedience to the law such that when dissenters say (and show) that they are employing this form of dissent, the general public will know more or less what to expect. Brown expresses this by arguing that it is important "to preserve civil disobedience as a tolerable, ritualized form of public protest in which law-breaking is minimal and for the most part formal." Cohen elaborates this idea with considerable subtlety in his article, but it is summarized rather well in the following short passage:

Going to jail does not "end" one's civil disobedience; on the contrary, it is a crucial part of that disobedience. It is crucial in characterizing the disobedience as being of a certain historically intelligible sort, and this, it is hoped, will have a special effect. For the fact that the disobedient displays his fidelity to law and strictly minimizes his deviation from it, as well as the fact that he is willing to suffer for his views, is meant to reassure and to move the majority.⁶

I hope it is clear enough now why it is important, from one perspective, to preserve what Brown calls "a ritualized form of public protest" and what Cohen sees as a form of protest especially appropriate for limited dissent to unjust laws or policies in an otherwise just constitutional democracy. I hope it is also clear that I have no objections to this view.

But now suppose that I say to Brown and Cohen: "Imagine two cases of disobedience in a more or less 'just' democratic society where all the circumstances and relevant factors are the same except that in-

dividual A is willing to accept the consequences but B is not. Are you saying that, because of that fact, A is justified in his disobedience (given that all the other conditions are met) while B is not?"

To this we must suppose that Cohen would answer: "He (B) may be justified, but he is not performing an act of justified civil disobedience."⁷ And I have no quarrel with this view. But Brown, as I have already suggested, maintains explicitly that an act of dissent in which the protestors do not willingly accept the (legal) consequences "is no less an act of civil disobedience, it is simply an unjustified one."

It is this latter view that I object to, but Brown's arguments support only Cohen's view and not his own much stronger view. Are there any better arguments for the latter view? I think that Professor Rawls attempts to defend some form of this view, though not successfully. Since Rawls's position is, superficially, much more like Cohen's than Brown's, I want first to determine to what extent Rawls wishes to defend some version of the latter view. I shall then try to see what sorts of arguments Rawls provides for the stronger, less tolerant position -- even if he does not explicitly endorse it -- and ask just how plausible they are.

2. Rawls's position with respect to paying the penalty is a curious one. It is clear from his remarks in section 55 of his book that a willingness to accept the legal consequences of one's conduct is an essential feature of Rawls's conception of justified civil disobedience (cf. especially pp. 366-67).⁸ Thus, he writes that although a particular law has been deliberately broken in cases of justified civil disobedience, "fidelity to law is expressed...by the willingness to accept the legal consequences of one's conduct" (p. 366). However, if we examine Rawls's list of the necessary conditions for justified civil disobedience (cf. section 57, pp. 371-374), we do not find the willingness requirement among them. Indeed, it is actually in the section of his book devoted to "The Definition of Civil Disobedience" (section 55) that Rawls discusses the question of "paying the penalty" or accepting the consequences.

This might mean that, like Professor Bedau in the article mentioned

in the preceding chapter, Rawls is merely defining a certain kind of disobedience to the law and saying that if the disobedient isn't willing to accept the consequences, then that's not "civil disobedience" in his sense. There is some evidence in the text that this is in fact what Rawls is doing. Thus, in one footnote he says explicitly that he is following "Bedau's definition of civil disobedience" (footnote 19, p. 364), and in another he says, apropos the question of accepting the consequences, that "there comes a point beyond which dissent ceases to be civil disobedience as defined here" (footnote 22, p. 366).

If Rawls, like Bedau, is making a willingness to accept the consequences part of his definition of civil disobedience, then I have no quarrel with him. Rawls, like Bedau, is perfectly free to define a certain form of dissent and say that when the appropriate features are absent dissent is not "civil disobedience" in his sense. But there is evidence in the text that Rawls wishes to make another, stronger claim, that he wishes to say that some acts of disobedience to the law which would otherwise be justified are not justified in the absence of a willingness to accept the legal consequences.

This is obviously quite different from saying that some acts of disobedience to the law which would otherwise be civil disobedience are not civil disobedience, in his sense, in the absence of a willingness to accept the consequences. Rawls obviously cannot make the former claim good simply "by definition." That is, the freedom to define civil disobedience in a particular fashion does not imply a similar freedom to stipulate when such conduct is and is not "justified." Arguments are required for claims about when civil disobedience, however defined, is or is not justified. What evidence is there for saying that, according to Rawls, the willingness requirement is relevant to the justification of civil disobedience as well as its characterization?

To begin with, it is worth noting that nothing like the "willingness requirement" is included in the actual definition of civil disobedience that Rawls gives in section 55; he writes:

I shall begin by defining civil disobedience

as a public, nonviolent, conscientious yet political act contrary to law usually done with the aim of bringing about a change in the law or policies of the government. (p. 364)

It may be objected that the willingness requirement is entailed by some feature(s) of this definition and in that sense is part of Rawls's definition of civil disobedience. I shall consider this claim in detail below when I discuss the relationship between nonviolence and a willingness to accept the legal consequences of one's actions. Nonetheless, it is significant that a willingness to pay the penalty is not mentioned explicitly in Rawls's definition of civil disobedience. What's more important, however, is that a whole class of illegal actions that would otherwise be justified are not justified, on Rawls's view, in the absence of a willingness to accept the legal consequences. It is this feature of Rawls's theory which makes it so controversial, and before proceeding I must show that this is indeed his view.

Rawls conceives of civil disobedience as occupying its own special position on a spectrum of possible reactions to legal authority, any of which can be appropriate depending on the circumstances of the actual society in which it occurs.⁹ At one end of this spectrum is straightforward obedience to the state's laws and at the other end is all-out rebellion. Between these extremes we find disobedience for the sake of testing a law's constitutionality, civil disobedience and conscientious refusal as Rawls conceives them, so-called "direct action", and what Rawls calls "militant resistance". The idea, very roughly, is that more radical means of protest are justified as violations of the principles of justice get progressively more serious and irremediable. Since this notion, which appears to me a very reasonable one in nearly every respect, occupies such an important place in my interpretation (below) of Rawls's views on the willingness requirement, I want to say a bit more about it here.

I wish to stay as close as possible to Rawls's actual way of putting things, but in the interests of clarity I shall have to simplify things a bit. Suppose we call the range of possible reactions to legal authority that Rawls alludes to the "reaction spectrum" in our

model of his theory.¹⁰ These possibilities range, as we have seen, from more or less complete obedience, to all-out rebellion. Let us say, quite arbitrarily and simply for the sake of clarity, that this spectrum ranges from "A" to "Z", where A represents more or less absolute compliance, where Z represents full-scale rebellion, and where the intermediate letter-variables represent the possible forms of dissent between these extremes (the forms mentioned in the preceding paragraphs: disobedience for the sake of testing a law's constitutionality, for example, direct action, militant resistance, and so forth).

Now beneath this spectrum we may imagine a correlative continuum describing the various circumstances in which a given form of dissent on the reaction spectrum is appropriate -- i.e., justifiable, from Rawls's point of view. We can call this the "circumstances continuum".¹¹ Let us say that the conceivable circumstances range from A', which would be a situation of perfect justice, to Z', which would be a situation of not only intolerable but monstrous injustice.

The circumstances of interest to the political theorist concerned with the possibility of justifiable forms of dissent short of all-out rebellion fall between these extremes, of course. Rawls is such a theorist and -- in terms of this very crude, simplistic model -- what he tries to show is that in a given social setting the appropriateness of a given form of dissent on the reaction spectrum is a function of the position that the society in question occupies on the circumstances continuum. As we proceed in the direction of Z' on the latter, more and more radical forms of dissent may be justifiable. That is, as we proceed toward Z' on the circumstances continuum, we proceed toward Z on the reaction spectrum. Needless to say, circumstances that would justify forms of dissent verging toward Z (all-out rebellion) also justify, a fortiori, any forms of dissent short of, or less radical than, those forms. But, and this is crucial to Rawls's theory as I understand it, until circumstances justify it (by changing in the direction of Z' on the circumstances continuum), a given form of dissent on the reaction spectrum cannot, justifiably, take a form of dissent closer to Z.

In short, and very simplistically, if circumstances in a given

society are described by G', say, on the circumstances continuum, then no form of dissent more radical than G, on the reaction spectrum, will be justified, although the possibilities short of G will still be available to the dissenters.

This view, especially as Rawls states it in section 55 of his book, is a very appealing one, and its simplicity does not detract from its theoretical power.¹² However, one feature of Rawls's view is quite controversial, it seems to me, and bears directly on the problem of the justification of civil disobedience and the willingness requirement. For suppose, as Rawls himself frequently does, that at times there will be circumstances such that while civil disobedience (of the sort described by Rawls) will be justifiable, no more radical means of dissent will as yet be justifiable. That is, suppose C represents "civil disobedience", in Rawls's sense, on the reaction spectrum, and D represents what Rawls calls "direct action". The latter form of dissent will not be justifiable, according to Rawls, until the "political sociology" of the situation, as he puts it, evolves from C' to D', for example.¹³

What disturbs me about Rawls's view can be illustrated by imagining a case similar to the one proposed to Cohen and Brown in section 1 above. Consider two individuals, A and B, who perform acts of disobedience of exactly the same sort, in the same circumstances, and for the same reasons. Suppose in addition that at the same time no further (or more "radical") steps would be justified. In other words, A and B are at the same point on the circumstances continuum and remain there for the purposes of our tale. Finally, suppose that both A and B meet all of Rawls's requirements for justified civil disobedience except for one: A is ready and willing to accept the legal consequences of his act but B is not.¹⁴

It may be that on Rawls's view B is not committing "civil disobedience"; this does not concern me at present. What does concern me is this: given the parameters of the situation we're imagining, and given Rawls's discussion of what is and is not justified in such situations (cf. especially p. 367), it seems clear that Rawls would say that A was

justified in what he did, under the circumstances, and B was not. And the reason would be that A met the willingness requirement while B did not (again, see p. 367). This is because for Rawls an unwillingness to accept the consequences is one step closer to Z, on the reaction spectrum, than an identical act of disobedience which exhibits a willingness to accept them.

What I wish to ask is simply this: what reasons does Rawls have, not for saying that A is, and B is not, performing an act of "civil disobedience", but for saying that A would be justified in what he did while B would not? I shall show that one answer to this question hinges on Rawls's conception of the notions of nonviolence and "fidelity to law." I shall also show that Rawls's reasoning in this connection is unsound.¹⁵

3. On Rawls's view both nonviolence and what he calls "fidelity to law" are essential to civil disobedience qua civil disobedience; that is, if these elements are not present, then we are not talking about civil disobedience but about some more radical alternative.¹⁶ I shall accept Rawls's characterization of civil disobedience as being by its nature non-violent and at the extreme limits of fidelity to law, although I shall subject the latter to a great deal of scrutiny below. What I want to determine is whether either or both of these conditions entails (or in some weaker sense "requires") that the disobedient must be willing to accept the legal consequences of his act.

Rawls has three arguments for saying that this is the case. He argues first that the nonviolence (as well as the "publicity") requirement entails the willingness requirement. He argues next that the willingness requirement is an indirect consequence of "fidelity to law," since fidelity to law entails nonviolence, which in turn entails the willingness requirement. He argues finally that the willingness requirement is an immediate consequence of fidelity to law; that is, that the latter entails the willingness requirement independently of the nonviolence requirement. This third argument is the most substantial since it can be rather easily shown that the nonviolence requirement does not entail the willingness requirement.

Rawls's definition of civil disobedience stipulates that to qualify as "civil" disobedience the illegal act must, among other things, be both public and nonviolent (p. 364). I shall subsequently refer to these as the requirements of "publicity" and "non-violence", and I shall accept them as noncontroversial elements of a reasonable conception (or definition) of civil disobedience. Does the willingness requirement follow from these definitional requirements?

At various points Rawls writes as though the public and nonviolent nature of civil disobedience entails that the disobedient must be willing to accept the legal consequences of his conduct (see, for example, section 57). If this is true, then such a willingness would be part of the "nature" (or definition) of civil disobedience and I would have no argument with Rawls. But it is not true. To be sure, there are certain ways of exhibiting an unwillingness to accept the consequences that would violate these requirements of publicity and nonviolence, and perhaps this is what has led Rawls to connect the willingness requirement with the requirements that the disobedience must be public and nonviolent. For example, if the unwillingness in question takes the form of forcibly resisting arrest, then the act is no longer nonviolent. Similarly, if this unwillingness is exhibited by the fact that the illegal act is done in private (in order to escape detection, say) then the act is obviously not "public" in the required way.

But we have agreed to eschew violence and act publicly when we wish to say that we are performing "civil" disobedience, and so these are not cases that Rawls can use to make his point. The question is: can an unwillingness to accept the consequences take any other form than resisting arrest, say, or attempting to keep one's criminal conduct secret? Surely it can. I can express my unwillingness to accept the consequences in writing (in a philosophy paper or a legal brief, for example); I can express it verbally (at the time of arrest, for example, or at my trial); or, most dramatically, I can express such an unwillingness simply by failing to show up for trial or by refusing to cooperate if I am forcibly brought to trial. None of these ways of repudiating the willingness requirement, if appropriately executed, need involve the disobedient in any

violation of the publicity and nonviolence requirements.

So the fact that civil disobedience must, by definition, be both public and nonviolent does not entail that the disobedient must be willing to accept the full legal consequences of his action. It is possible to meet the former requirements while denouncing the latter, and it seems to be mistaken to suppose otherwise, as Rawls does. In fairness to Rawls, however, we must grant that considerations like the preceding force us to clarify what we have thus far left rather inchoate: namely, just what it means to repudiate the willingness requirement. We see now that we cannot express our unwillingness to pay the penalty by means of force or secrecy; we cannot forcibly resist arrest, for example, or try to escape detection and arrest by acting secretly. But the modes available to us for expressing an unwillingness to accept the consequences are not exhausted by these alternatives, as I tried to point out above.

More significantly, perhaps, we have isolated a far more important issue: the crucial issue in the repudiation of the willingness requirement is not the practical question of whether or not we can manage to avoid punishment (without violence or secrecy), but the question of whether or not we are going to grant, in principle, that the state retains a moral right to punish us in cases of the sort we're discussing. Thus, I shall henceforth construe "accepting" the willingness requirement as accepting the claim that the state retains its right to punish (and exact legal consequences generally) in the cases in question, cases where the illegal act is, ex hypothesi, "justified". And I shall construe "rejecting" the willingness requirement as denying this claim. I must postpone until Chapter VI the question of what means the disobedient can legitimately resort to -- other than violence or secrecy -- if he rejects the willingness requirement in principle.

4. I should like to turn now to the question of whether "fidelity to law" entails a willingness to accept the consequences, but first I must show how Rawls persists in thinking of an unwillingness to accept the consequences as necessarily involving violence, even when he is discussing "fidelity to law" as a rationale for the willingness requirement. Consider

the following remarks:

Civil disobedience is nonviolent for another reason. It expresses disobedience to law within the limits of fidelity to law, although it is at the outer edge thereof. The law is broken but fidelity to law is expressed by the public and nonviolent nature of the act, by the willingness to accept the legal consequences of one's conduct (p. 366).

Both the underlying logic as well as the grammar of this quotation are extremely suggestive. Logically, the structure of the argument is this:

- (1) Civil disobedience is justified (*qua* civil disobedience) only if it is disobedience "within the limits of fidelity to law".
- (2) Disobedience is within these limits only if it is public and nonviolent.
- (3) Civil disobedience would not be public and nonviolent if the disobedient were not willing to accept the full legal consequences of his illegal act.
- (4) Therefore, civil disobedience is justified only if the disobedient is willing to accept the full legal consequences of his action.

I propose to accept premises (1) and (2) without any argument for now. Since along with premise (3) they entail (4), it is this third premise that concerns me here. That Rawls is making a claim like that suggested in (3) is implied both by the structure of his argument as a whole and by the grammar of the second sentence in the preceding quotation: Rawls does not say that fidelity to law is expressed by the public and nonviolent nature of the act as well as by the willingness to accept the consequences; he writes as though a willingness to accept the consequences is somehow synonymous or conceptually connected with an acceptance of the fact that the act must be public and nonviolent. The crucial clauses are connected not by the particle 'and' but by a comma.

This is not just a verbal quibble. The grammar of Rawls's assertion suggests that for him we shall not be meeting the nonviolence requirement if we are not willing to accept the consequences of our action. But we have already seen that this is not necessarily true. That is, we have seen that the willingness requirement is at least not obviously en-

tailed by the requirements of publicity and nonviolence.

Rawls's first argument from the notion of fidelity to law fails, therefore, for the same reason that the argument from publicity and nonviolence fails: the latter notions do not entail a willingness to pay the penalty. Rawls has another argument based on the notion of fidelity to law, however. This is an argument which (in effect) concedes that an unwillingness to pay the penalty can be both public and nonviolent, but which purports to show that nonetheless such an unwillingness is not consistent with the principles of fidelity to law, even when that unwillingness is expressed both publicly and nonviolently.

The new argument is roughly this. Civil disobedience is distinguished from "militant action" (and other more radical forms of dissent) by the fact that the former, and not the latter, is within the bounds of fidelity to law.¹⁷ This is important because there are times when disobedience to the law is justified but when it must, for reasons to be explained, remain within the limits of fidelity to law in order to be justified. Let us summarize this by saying that civil disobedience is justified, qua civil disobedience, only if it is disobedience within the limits of fidelity to law. With this claim as its first premiss, the new argument continues with the following claim as its second premiss: civil disobedience is not disobedience within the limits of fidelity to law unless the disobedient is willing to accept all the legal consequences of his conduct. The conclusion of course is that civil disobedience is justified only if the disobedient is willing to accept all the legal consequences of his conduct.

I want to show that the second premiss in this new argument is false and that an unwillingness to pay the penalty is not inconsistent with a commitment to fidelity to law. In order to do this I must first explain briefly what fidelity to law is and describe the role it plays in Rawls's theory of civil disobedience. I am particularly anxious to show why civil disobedience must be within the bounds of fidelity to law, according to Rawls. Then I shall show that an unwillingness to pay the penalty, provided that it is unconcealed and nonviolent, does not necessarily take one beyond those bounds.

Rawls assumes, quite correctly I believe, that societies and the institutions that comprise them are not always absolutely just or unjust but are frequently only more or less so. In terms of the figure sketched in section 2 above, we can imagine that there is a continuum describing (very roughly) the range of possibilities with respect to the overall justice or injustice of various human societies. This would be a continuum ranging from a more or less perfectly just society at one extreme to a more or less totally corrupt society at the other extreme. In the middle will be societies that vary in one way or another from these extremes. In a totally corrupt or unjust society men are presumably justified in all-out rebellion. On the other hand, Rawls argues that in a reasonably just society¹⁸ men not only have a moral duty to obey just laws and policies, but sometimes even have a moral duty to obey unjust legislation as well.

I shall have a great deal to say about the latter claims in Chapter IV below. For the present I simply want to point out, what Rawls himself makes quite clear, that it is a corollary of this view that the extent and nature of justifiable dissent (within a given society) depends upon roughly where that society is located on the "circumstances continuum" described above. Thus, it is conceivable (indeed, highly probable, according to Rawls) that a given society will depart far enough from the ideals of justice to warrant moderate dissent but not so far from these ideals as to warrant an unqualified repudiation of the society as a whole and the moral authority of its legal system generally.¹⁹ That is to say, a given society may be sufficiently corrupt or unjust to forfeit its claims to moral authority with respect to certain unjust laws or policies without at the same time forfeiting its moral authority generally. Hence, a reasonably just society retains its overall moral authority and legitimacy -- its right to command our allegiance and cooperation -- even when certain of its laws and policies have lost their moral force.²⁰

The point of Rawls's theory of civil disobedience is to show how this particular form of dissent can be justified in a reasonably just society consistent with a recognition of the prima facie moral duty that one has to abide by duly enacted laws and policies in such a society, even when they are unjust. It is extremely important, therefore, to bear in mind that the

theory of civil disobedience is designed to speak to problems that arise only at a very special point on the so-called "circumstances continuum". We are imagining a situation in which laws of substantial injustice have been passed, so that civil disobedience is justifiable, but in which injustice is not so pervasive as to justify any steps beyond civil disobedience. In particular, we would not be justified, in this situation, in taking any steps that would be destructive of the society's overall stability, as Rawls sometimes puts it.

In light of these rather sketchy remarks, it is fairly easy to say roughly what fidelity to law is and why it occupies so prominent a place in Rawls's theory of civil disobedience. We are expressing fidelity to law in our dissent when we are careful not to do anything that will undermine the society's long-range social and political stability, stability which Rawls believes can be attained and preserved only through "the rule of law".²¹ Conversely, we are not expressing this fidelity when we act in ways likely to jeopardize that stability and the effects of the rule of law generally. "Fidelity to the task of upholding what we recognize as a reasonably just society and the rule of law that sustains it" would be a more apt characterization of this notion. But it would also be more awkward, so I shall stick to "fidelity to law".²²

The importance of staying within the limits of fidelity to law, in cases of the sort we are imagining, should also be obvious if we accept Rawls's general view. For we have both a natural duty (and sometimes an obligation) on that view "to support and to further just institutions" (p. 334).²³ We are, to be sure, dealing with a situation in which we are confronted with substantial injustice, a situation in which civil disobedience is justified. But we are also dealing, ex hypothesi, with a situation in which the society and its laws and institutions are, on balance, still reasonably just. So we have a natural duty to see to it that our actions remain within the boundaries of fidelity to law, as defined above. It is not an abstract fidelity to an abstract conception of "law" that is at stake here, but a very concrete commitment to the legal and constitutional structure of a given society.

We can now return to the main thread. It is easy to see, from what

has just been said, why the notion of fidelity to law occupies such an important place in Rawls's theory of justified civil disobedience and, in particular, why he believes that this notion precludes violence and organized militant resistance. The theory of civil disobedience is relevant when some valid law or duly enacted policy seems to exceed the bounds of tolerable injustice. But as long as we are dealing, by assumption, with a society that is reasonably just as a whole, despite occasional lapses, then we must acknowledge that there are serious limits on how far a citizen may go in protesting the injustice of a particular law or policy. Very roughly, these limits are those that define when we are still acting in accordance with the spirit if not the letter of the law. Rawls believes that public, nonviolent dissent aimed at addressing the sense of justice of the majority is within the spirit of the rule of law, even when it transgresses the letter of a particular law or policy. But to engage in violent actions, for example, in an attempt to right social wrongs by force rather than persuasion, is clearly outside the limits of fidelity to law and is a violation of the spirit as well as the letter of law.

This is an unhappily abstract way of putting Rawls's point, but unfortunately he himself is neither clearer nor more concrete. The main point is obvious enough, however. Under certain circumstances illegal dissent is justified in a more or less just democratic society, but it must be limited dissent, limited by our commitment to the system as a whole and our responsibility not to do anything likely to undermine it. Suppose that violence is the sort of thing likely to undermine the mutual trust and social stability that are required to manage a reasonably just constitutional democracy. We shall then say that if nothing more radical than civil disobedience is justified at a given time in a given society, then that disobedience must be nonviolent. This is because we've conceded that nonviolence is inconsistent with fidelity to law. However, Rawls also claims that an unwillingness to pay the penalty for civil disobedience is inconsistent with fidelity to law, and hence that if nothing more radical than civil disobedience is justified under certain circumstances, then one is not justified in seeking to escape the consequences of one's illegal act, even if the act itself is justified. Why is this

so, according to Rawls?

The short answer is to say that in the absence of a willingness to pay the penalty the disobedient's conduct would be subversive of the stability of the society in question and of the rule of law generally; but of course this is exactly what needs to be argued, since this claim is at best controversial. Moreover, I do not think that Rawls wants (at least explicitly) to urge a simple "subversion" claim, like that advanced by Socrates and the Laws in the Crito for example (but see Chapter III below). Rather, I think Rawls is making a much more sophisticated claim. Roughly, his claim is that an unwillingness to pay the penalty is subversive of the aims and principles of democratic constitutionalism. This notion must now be explained.

Consider the following quotation from Chapter VI of Rawls's book:

...to the legal forms of constitutionalism one may adjoin certain modes of illegal protest that do not violate the aims of a democratic constitution in view of the principles by which such dissent is regulated (p. 386).

This is somewhat opaque, but it contains the kernel of Rawls's most important (explicit) argument for the willingness requirement. Dissent is acceptable, in a reasonably just constitutional democracy, provided that it does not violate the aims of the just constitution. We ensure that it will not do so by regulating such dissent by certain principles. One of these principles is the willingness requirement. Now according to Rawls the aims of a constitutional democracy are, very roughly, the creation of just and efficient social institutions. If we interpret the willingness requirement as one of the principles by which dissent is regulated consistent with these aims, then obviously the argument for the willingness requirement must involve the claim(s) that to eliminate it would lead to less just and/or less efficient institutions. In the remarks that follow I shall explicate the arguments for these claims and show that they are not sound. That is, I shall show that a "principled" (i.e., public and nonviolent) repudiation of the willingness requirement is consistent with fidelity to law as Rawls conceives it. To do this I shall have to show what tempts Rawls to say that this is not the case.

At first blush it would seem that we ought to separate the question of whether refusing to pay the penalty would simply be unjust from the question of whether such a refusal would be detrimental to the efficiency of the society's institutions. That is, it seems that we ought to separate the question of whether there is a more or less "deontological" rationale for the willingness requirement from the question of whether there is a more or less "teleological" rationale for that requirement. It would be convenient to be able to do this, and I shall attempt to do so below. But first a caveat: Rawls and Professor Charles Fried, the principal authors we shall be considering, do not always keep these considerations separate. Hence, it is not always clear whether they are claiming that refusing to accept the consequences is unacceptable because it is in itself "unfair" or because it leads to social instability or both.

Consider the claim that one is violating the principle of fair play if one disobeys a valid law and then refuses to pay the penalty. Rawls has argued in another context that we have an obligation, based on the principle of fair play, to obey unjust laws as long as the constitution itself is reasonably just and the laws in question do not exceed certain limits of injustice.²⁴ One might be tempted to argue that we have a similar obligation to pay the penalty after breaking a law that exceeded the tolerable limits of injustice. But it's difficult to see the analogy between these two cases.²⁵ We are dealing, ex hypothesi, with situations of grave injustices. Rawls himself concedes in a footnote that "certainly one does not accept the punishment as right, that is, as deserved for an unjustified act" (p. 366). But he continues: "Rather, one is willing to undergo the legal consequences for the sake of fidelity to law, which is a different matter" (ibid.). Rawls's reasoning here is remarkably obscure. If the punishment is neither "right" nor "deserved for an unjustified act", how can punishing and requiring acceptance of punishment be right? Rawls makes a valiant effort to answer this question, but his answer is clearly not in terms of the deontological notion of "fair play" but in terms of the likely social consequences of not accepting (and requiring acceptance of) the legal consequences of illegal action.

Rawls has argued earlier that the danger in allowing indiscriminate

disobedience to the law, even on conscientious grounds, is a greater evil than the evil of requiring people to obey at least some unjust laws. Similarly, it might be argued, the instability of a system where we "let people off" if their reasons were convincing would be a greater evil than the evil of punishing people who were ex hypothesi justified in disobeying the law. Thus, Rawls writes:

No doubt it is possible to imagine a legal system in which conscientious belief that the law is unjust is accepted as a defense for noncompliance. Men of great honesty with full confidence in one another might make such a system work. But as things are, such a scheme would presumably be unstable even in a state of near justice (p. 367).

This passage involves a serious confusion on Rawls's part, I think. He says that, as things are, "...a legal system in which conscientious belief that the law is unjust is accepted as a defense for noncompliance... would presumably be unstable...." (p. 367) Let us suppose that this is so. What bearing does this fact have on the theoretical question of whether it is reasonable to expect individuals to accept the ordinary legal consequences for disobedience, when they are justified, ex hypothesi, in that disobedience? Rawls's remarks are relevant to the issue of whether a sincere belief that the law is unjust is a sufficient condition for noncompliance. I agree with him that it is not. Such a belief does not even ensure that the act itself is justified (independent of its illegality), much less that noncompliance (in the form of civil disobedience) is justified. But Rawls also believes that noncompliance can be justified; surely this is his central point in the latter half of Chapter VI of his book.

The question I am raising, then, is simply this: when civil disobedience is justified on the grounds suggested by Rawls, what reason is there to suppose that we ought nonetheless to punish the disobedients? It is no answer to this question to say that sincere belief does not by itself justify noncompliance. Of course it doesn't. But when we get clear about what does justify noncompliance -- and I take it that this is what section 57 in Rawls's book is all about -- then presumably we have said what more is needed to justify civil disobedience. As long as we

have no clear conception of what justified civil disobedience is, then perhaps considerations of "stability" make it necessary to allow civil disobedience only if the disobedient is willing to accept the legal consequences of his action. But if Rawls has achieved anything in Chapter VI, he has managed to help us get clear about just when civil disobedience is justified. What he doesn't appear to recognize is that this raises a whole new series of questions -- the most important of which is just the one I have been belaboring: if a man is ex hypothesi justified in disobeying the law, then how can society be justified in punishing him?

5. None of Rawls's arguments in Chapter VI provides an adequate answer to this question, as I hope I have shown. This, of course, does not show that there is no good reason for making the willingness requirement a part of the theory of justified civil disobedience; it shows only that Rawls has not provided such reasons. In Chapter V I shall elaborate a general argument that in fact any theory of civil disobedience that justifies that disobedience on the basis of a "contractarian" theory of political obligation is incoherent if it also includes a requirement that the disobedient must accept all the legal consequences of his conduct. Before turning to this enterprise, however, I should like to look at Professor Charles Fried's recent attempt to support the willingness requirement with principles borrowed from Rawls.²⁶

Fried's argument is based on an interpretation of Rawls's principle of fair play. He has shown in an earlier argument that it is this principle that makes it wrong for some individual to violate a just income-tax law on the grounds that his delinquency will obviously help him and in no way hinder the general welfare. Even if the latter claims are true, Fried argues, the tax-evader is acting unjustly because he is taking advantage of benefits that can be secured only by general observance by others of the very law he is violating. Fried argues that the same charge can be pressed against the civil disobedient, albeit "at another level of generality":

...he [the civil disobedient] seeks to benefit from

an institution which depends on the sacrifices of others without contributing a like sacrifice. Only here the institution and the sacrifice are respectively, the legal system as a whole and the readiness of individuals within that system to abide by the principle of "institutional settlement", that is, the readiness of each to forego pressing his claim -- even when he considers his claims justified -- when by some fair procedure the issue has been determined against him.²⁷

Fried makes two distinct points about the violation of "the principle of institutional settlement", although he writes as though he is unaware of just how distinct they are. The first is this:

The demonstrators in violating this principle run the risk that if they are successful in procuring a change in the law those persons who believe that the old situation was justified and that the new unjustifiably prejudices their interests will feel entitled to resist the "remedial" legislation.²⁸

While this observation may be sound as a comment on the psychological dynamics of dissent, surely Fried would be the first to admit that feeling entitled to act in a certain way must be firmly distinguished from being entitled to act that way. Hence, whatever conclusions the demonstrators draw from Fried's remarks so far will be what we have called tactical or strategic ones: they have to take the likely consequences of their actions into account. If the only way to prevent a vicious cycle of disobedience and counter-disobedience is to accept the legal consequences of their initial disobedience, then perhaps this is what they ought to do on balance.

But Fried wants (and needs) to make a much stronger claim than this. Apropos the feeling that one is entitled to resist (illegally) "remedial" legislation that was brought about through civil disobedience, Fried writes:

...this feeling would not be wholly unwarranted, as the demonstrators' unwillingness to abide by the principle of institutional settlement does weaken their moral claim on the fidelity of others to that principle.²⁹

Fried offers no argument for the crucial moral claim embedded in this

remark, but he does offer a "solution" for the practical problem of cyclic disobedience that he thinks arises in such cases -- viz., a willing acceptance of the legal consequences of one's illegal act. The civil disobedient, "in showing his readiness to suffer the penalty for disobedience," is, according to Fried, thereby "affirming the value of law in general, while at the same time disobeying it."³⁰ Perhaps it is true that in paying the penalty the disobedient is in some sense affirming the value of law in general, although I should like to have seen some argument for this claim. But surely Fried is mistaken in asserting that the civil disobedient has weakened his moral claim on the fidelity of others to the principle of institutional settlement. For built into Rawls's theory of civil disobedience is the admission that in certain cases one has a moral right to disobey the law. How can the exercise of this right deprive him of his otherwise legitimate expectation that others will abide by the rule of law?

There is, of course, one last move open to Rawls and Fried here. They can agree that as long as the disobedience takes place within the limits of "fidelity to law" the disobedient has not lost his moral claim on others to stay within those limits. They can then stipulate that unless he is willing to accept the consequences, the disobedient is not acting within the limits of "fidelity to law". Unfortunately, this move guarantees inclusion of the willingness requirement at the expense of begging the question. Why does the disobedient have to accept the consequences to stay within these limits? Far from having forfeited his moral claim by his disobedience, on Rawls's theory he is reminding others of the principles underlying the just constitution and pointing out that they have been violated. It may be prudent for him to accept the consequences, but as yet we have seen no convincing argument that he must, in principle, do so.

I said at the beginning of this chapter that I would be considering only Rawls's "explicit" arguments for the willingness requirement. If these arguments do not seem very impressive, I hope this is not because I have misinterpreted Rawls's remarks with respect to paying the penalty. I have simply gone carefully through the text, explicating several obscure

remarks and amplifying them with an argument from Fried that Rawls himself seems to consider conclusive. However, I do not feel that Rawls himself actually does justice to the willingness requirement. As implausible as this requirement is at first glance, and as implausible as it remains after we have considered Rawls's explicit arguments in favor of it, I think there is a far more powerful argument for the willingness requirement than any argument Rawls explicitly advances in behalf of it. It is an argument that is entirely within the spirit of the Rawlsian enterprise, however, and it is perhaps fair to say that it is implicit in Rawls's work. In the next three chapters I should like to develop this argument, partly because I believe it suggests why reasonable men of good will might claim that they were right in punishing even justified disobedience to the law and also because I am anxious to make such implicit arguments explicit and show that in the long run they too are unconvincing.

CHAPTER III

THE "LEGALIST" TRADITION AND SOCRATES' ARGUMENTS IN THE CRITO

1. The whole course of our discussion thus far may seem somewhat dissatisfying. Surely there are better arguments for the willingness requirement than any we have so far considered. Or perhaps more can be said in favor of some of the arguments I have sketched above than has yet been said. I think that in fact there is another argument for the willingness requirement that is both more powerful and more interesting than any considered above. But this argument is at best only implicit in some of Rawls's work on political obligation generally; and perhaps it would be most apt simply to say that we can construct an impressive argument for the willingness requirement by analogy to another argument that appears explicitly in Chapter VI of Rawls's book and in a number of his earlier publications.

I am alluding to Rawls's argument for the claim that, at least in a reasonably just constitutional democracy, we have a prima facie moral duty (and sometimes a moral obligation) to obey the law. What is particularly interesting about this argument is that it purports to show how it happens that sometimes we have an "actual" or "on balance" moral duty to obey even unjust laws. Rawls's argument for the latter claim is of interest in its own right, as I shall show below. It is also interesting, however, because in light of it we can construct the "argument by analogy" for the willingness requirement: the argument that just as we are sometimes morally required to comply with unjust legislation, so too, under certain circumstances we are morally required to accept the results of fair judicial procedures, even when this means accepting the legal consequences for disobedience that was itself morally justified. Before any of this will make much sense, something must be said about the tradition in political theory in which Rawls seems to stand.

The intellectual atmosphere in which most recent thinkers have worked can be called an atmosphere of qualified (and sometimes quite extreme) "legalism", an atmosphere in which it is taken for granted that the moral claims of civil law are very heavy ones and that the burden of proof is on those who would disobey even an admittedly unjust (or unfair) law rather than on those who would enforce such a law.¹ Professor Hugo Bedau may have had this contemporary moral and political bias in mind when he complained recently that one of the most disturbing features of recent writing about civil disobedience is its uncritical assumption that the moral burden in questions of obedience to the law and enforcement of it is on the individual in his disobedience rather than on society in its enforcement.² That is, it is taken for granted that we have a presumptive moral obligation to obey the law even when it is somewhat unjust; the problem is conceived as that of showing when this presumptive (or prima facie) moral obligation breaks down.³

Professor Bedau made his complaint apropos a very recent book on the problem of civil disobedience.⁴ But it should have been obvious in the early 1960s, when theoretical interest in the justification of civil disobedience was revived (at least in the United States) by the civil rights movement, that the whole cast of the discussion was to embody the assumption in question.⁵ In one of the most significant articles of that period, Professor Richard Wasserstrom wrote that the view he was "most concerned to show to be false" was the view "that one has an absolute obligation to obey the law."⁶ And yet it would be very odd to find that any respectable philosopher or political theorist had ever actually held this view, sans phrase, since it seems to entail that not even the most oppressed mass of people imaginable would be justified in resisting the most corrupt tyrant conceivable as long as the latter's edicts were embodied in valid law.⁷ Why, then, should Wasserstrom be so concerned about this view? Apparently he is not, after all. He is, despite what he says, actually concerned about a qualified, but still very extreme, version of legalism.

The view that Wasserstrom is in fact concerned with is the view that we have an absolute obligation to obey the law whenever we have any moral obligation to obey the law at all. Wasserstrom interprets western political

theory as embodying the claim that we are never justified in disobeying the law in a given society unless all-out rebellion is justified in that society. This qualification removes the cases which make the view Wasserstrom is so concerned about totally untenable, but the view we are left with is still a very strong one: if we are not justified in all-out rebellion, we are not morally justified in any sort of disobedience or illegal dissent short of rebellion.⁸

Notice that we are still left with Professor Bedau's question: how did the cards ever get stacked so heavily against the moral right of the individual to resist injustice short of the gross injustice that justifies rebellion? One looks in vain for arguments in the works of the philosophers Wasserstrom cites as exponents of this view, since it seems that they did not in fact expound it.⁹ Indeed, it would be odd if they did. Wasserstrom attributes this view, for example, to Hume. Hume does indeed argue for a presumptive moral obligation to obey the law, but he does not argue that this obligation is absolutely binding up to the point when revolution is justified. And in order to do so, he would have to make a very curious empirical assumption, one which to my knowledge he never makes. The point is this. Hume's argument for the obligation to obey the law is a teleological or a "utilitarian" one: we are morally required to obey the law because in general the consequences of doing so are better than the consequences of not doing so.¹⁰ Presumably, a situation where revolution would be justified is a situation where the over-all negative utility of continued obedience obviously outweighs the over-all positive utility of doing so, to put it crudely. But in order for Hume to be defending the normative claim that no sort of disobedience short of rebellion can ever be morally justifiable, he would have to be defending the empirical claim that the over-all positive utility of disobedience to the law will as a matter of fact be greater than the over-all negative utility of such disobedience only when conditions are bad enough to justify rebellion.

Perhaps it is sufficient to point out that this is an empirical claim that Hume never makes, but it is hard to resist pointing out what an implausible claim it is as well.¹¹ Contemporary utilitarian defenses

explicitly recognize this when they make the conditions under which disobedience to the law is justified a direct function of the long-range utility of such disobedience and when they point out that positive utility can outweigh negative utility in all sorts of situations short of revolutionary ones.¹²

The preceding point can be generalized in the following way. Suppose a theorist appeals to moral principles like "justice" or "utility" as the basis for our obligation to obey the law. Since the author of this doctrine is appealing (ex hypothesi) to these principles in accounting for our obligation to obey the law, it seems always to be at least a theoretical possibility that these same moral principles can be invoked as a justification for disobeying the law. Now if considerations of justice or utility can justify all-out rebellion, why can't these same moral principles justify something short of revolution? It seems that any philosopher who bases the obligation to obey the law on some moral principle like justice or utility must face this general question; that is, it seems that his principles will cut both ways (for and against disobedience to the law) depending on the circumstances. Why should they favor disobedience only in the worst imaginable circumstances?¹³

2. To be fair to Wasserstrom, it should be conceded that at least one philosopher he mentions, although not in the context of his actual historical claim, does appear to hold the view that Wasserstrom, along with Professor Marshall Cohen, regards as characteristic of traditional democratic theory in western political thought. I am alluding, of course, to Socrates and the position he defends in the Crito. It is interesting that Wasserstrom says that his own aim is to determine the nature and extent of our obligation to obey the law and interesting that he goes on to say that this problem "received at the hands of Socrates...a sustained theoretical analysis and resolution" in the Crito.¹⁴ For one thing, Wasserstrom himself never gets around to telling us just what the nature, much less the extent, of our obligation to obey the law is. At best, he has managed to refute two extremely implausible views: (1) the view that civil disobedience is never justified in a more or less just society;

and (2) the view that civil disobedience is never justified in a more or less just democracy. But Wasserstrom's remarks are interesting for another reason: the one clear case he gives of a philosopher who actually seems to have held one or the other of these views is that of Socrates in the Crito; and as we shall see below, Socrates' arguments are extremely unimpressive.

A great many discussions of civil disobedience begin by paying homage to Socrates' eloquent defense of some form of extreme legalism in the Crito. It will be worthwhile, for a number of reasons, to recall the actual arguments that Socrates uses there. For one thing, these arguments are, as I have said, strikingly unimpressive ones, as we can see simply by departing from what is customary in this context and actually laying out the arguments Socrates relies upon in the dialogue. But this enterprise, as surprising as its results may be, would be of little more than historical (and perhaps iconoclastic) interest, if it were not for another striking feature of Socrates' arguments. Despite the fact that they are, as they stand, rather bad arguments, they anticipate -- in an almost uncanny manner, as I shall show -- some very powerful arguments that have recently been advanced by Professor Rawls both with respect to our putative obligation to obey the law generally and our putative obligation to accept the legal consequences of disobedience under certain conditions, when the moral obligation to obey a given law breaks down.

Rawls's arguments are more powerful than those of Socrates not because the former are developed more carefully and explicitly (they're not), but because they aim to prove a more modest conclusion: not that we have an absolute obligation to obey the law in a reasonably just society but that we have a prima facie obligation to obey the law which generates, sometimes, an "actual" or "on balance" obligation to obey virtually all just laws in such a society and even a great many unjust laws as well. I am primarily interested in Socrates' arguments because of the curious way they anticipate Rawls's arguments and because their failure is instructive in ways I shall indicate. I am interested in Rawls's

arguments, or rather, in reconstructing arguments for his inchoate enthymemes and argument-sketches, for a number of reasons.

To begin with, Rawls provides us with at least a sketch of a possible answer to Professor Bedau's demand for arguments that show that there is even a presumption, morally speaking, in favor of obeying the law as opposed to acting by one's private moral lights. Secondly, Rawls takes up, to some extent, the burden of telling us what it is to show that certain conduct is or is not "morally justified," whatever its legal status, before arguing that civil disobedience is in fact morally justifiable. Thirdly, in attempting the previous task, Rawls puts his own moral principles on the table as it were and thus enables us to determine for ourselves just how plausible the moral and political bias that Bedau discerns in other theorizing actually is, at least insofar as Rawls's principles are anything like those of other less articulate political theorists. Finally, and perhaps most importantly for the purposes of the following chapters, Rawls's arguments for restricted or qualified legalism with respect to legislative decisions can be used to formulate analogous arguments for the view that I am ultimately most interested in: the view that in a reasonably just constitutional democracy, even when the prima facie obligation to obey the law breaks down and we are morally justified in civil disobedience, we are nevertheless morally required to accept the full legal consequences of our illegal actions. I shall eventually show that the latter view is best characterized as extreme legalism with respect to judicial decisions in a reasonably just society and can be contrasted with Rawls's restricted or qualified legalism with respect to legislative decisions in such a society. Both views are part of that "legalistic" tradition in political theory that goes back at least as far as Socrates and that Professor Bedau has criticized recent thinkers for accepting too uncritically. The point of this and the following chapter is twofold. For one thing, the arguments examined will familiarize us with a point of view (concerning the moral force of civil law) which makes the contemporary emphasis on paying the penalty quite understandable. Secondly, and much more importantly, these arguments will enable us to construct an analogous (and, I think, very

powerful) argument for the willingness requirement as an essential part of any theory of justified civil disobedience.¹⁵

3. It is curious, on the one hand, that so many discussions of the justifiability of civil disobedience should allude to Socrates' position in the Crito, since Socrates suggests at several places in the dialogue that the question he is addressing is not the general one of the justifiability of civil disobedience but the much more specific one of the moral propriety of accepting the consequences of a fair trial, even when the verdict is "unjust," as Socrates puts it, or simply mistaken, as we should say.¹⁶ It is easy, on the other hand, to understand why so many readers should interpret the Crito as a tract on political obligation generally, and civil disobedience in particular, since the arguments that Socrates uses to convince Crito that it would be "unjust" or morally wrong to try to escape from prison seem to imply that it would be wrong to disobey the law under any circumstances, at least in a more or less just democratic society.

If we distinguish legislative decisions (which are embodied in laws enacted by duly constituted legislative bodies) from judicial decisions (which are embodied in the decisions of duly constituted judicial bodies or courts), then it is obvious that Socrates' explicit intent in the Crito is to defend some form of extreme legalism with respect to judicial decisions, at least in more or less just societies. Socrates indicates as much when he structures the inquiry by telling Crito that the issue they must face is "whether it is just or not for me to try to escape from prison, without the consent of the Athenians,"¹⁷ and when he insists that they "have nothing to consider but the question which I asked just now -- namely -- shall we be acting justly...if we ourselves take our respective parts in my escape? Or shall we in truth be acting unjustly?"¹⁸ It is quite clear, then, that the question Socrates explicitly raises is the narrow question about our obligation to comply with fair judicial procedures in a just society rather than the much larger question about our obligation to obey the law generally in such societies. Despite the fact that Socrates himself tends to

conflate these different questions in the course of the dialogue, it is important to keep them separate in our own minds, for a number of reasons.

To begin with, while extreme legalism with respect to legislative decisions (in a reasonably just democracy) certainly seems to entail extreme legalism with respect to judicial decisions (in such a society), the converse is not the case. That is, it is possible to contend that we have an absolute obligation to abide by fair judicial procedures in a reasonably just society without contending that we have an absolute obligation to abide by all laws or legislative decisions in such a society. Indeed, in Chapter IV I shall show that this inverted legalism is something like the view of Professor Rawls in A Theory of Justice. And as I shall show in this chapter, Socrates seems to anticipate Rawls's view in the Crito.

Another reason for keeping these two views about the nature of legal obligation separate has to do both with the arguments for the willingness requirement considered above and especially with the argument evaluated in Chapters V and VI below. All the recent thinkers who, in elaborating theories of the justifiability of civil disobedience, have done so much to refute extreme legalism with respect to legislative decisions, are at the same time anxious to insist on extreme legalism with respect to judicial decisions, at least in reasonably just democratic societies. The latter view, of course, is embodied in what I call the willingness requirement that is an essential part of so many recent theories of civil disobedience. The explicit arguments for the willingness requirement in the recent literature on civil disobedience are no better than Socrates' arguments in the Crito for the more general form of legal absolutism (extreme legalism with respect to legislative decisions), to which I now turn. In the next two chapters I shall show that the best argument for the willingness requirement (or extreme legalism with respect to judicial decisions) is one that can be constructed by analogy to Rawls's argument for restricted legalism with respect to legislative decisions. I should like to proceed to the latter argument via a brief discussion of Socrates' arguments in the Crito.

4. There are two preliminary points that it is quite important to bear in mind in any discussion of the Crito. To begin with, the issue of whether to try to escape from prison is a moral issue for Socrates. He says that the question he is about to examine in the Crito is a question of "...justice and injustice...the base and the honorable...the good and evil."¹⁹ In arguing the urgency of a correct answer to the question Socrates insists that "...we should set the highest value, not on living, but on living well," and he goes on to say that "living well and honorably and justly mean the same thing."²⁰ I shall explain why this assumption is so important in Section 5 below.

It is also important, by the way of preliminaries, to note that the argument against escape (and in favor of accepting the consequences of his trial) proceeds on the assumption that Socrates was in fact not guilty. Thus, Socrates places a great deal of emphasis on the moral principle that it is wrong to repay injustice with injustice. The clear implication, throughout the dialogue, is that the first injustice was the injustice of the verdict against Socrates. Indeed, when Socrates personifies the laws of Athens and has them deliver the argument against attempting to escape, they themselves acknowledge that if he goes to his death Socrates will have suffered a grave injustice -- the injustice of a mistaken verdict; but they emphasize that in this respect he is "a victim of the injustice, not of the laws, but of men."²¹ Finally, Socrates repeatedly finds it necessary to remind Crito that his own innocence is irrelevant to the question at hand. It is important that, in Socrates' view, his innocence is irrelevant; but it is also important that, on his view, he is innocent.

The assumption that he is in fact innocent is of the first importance in understanding both the precise thesis Socrates is defending as well as the various arguments against escape that appear in the dialogue. Indeed, the dialogue would not be of much philosophical interest if the arguments were not based on this assumption. For these arguments are designed to show not merely that we must invariably accept the consequences of disobedience to the law, but that we must accept the consequences of a fair trial even when we have been wrongly convicted or when

we have broken the law but were justified in doing so. This thesis is particularly startling when we realize that Socrates is saying that he has a moral obligation to let the state execute him for crimes he did not commit. Why, given the fact that the state has "wronged" or "injured" him by unjustly finding him guilty and sentencing him to death, would it be wrong or unjust for Socrates himself to try to escape?

Despite a great deal of subsequent unclarity, Socrates' answers to this question are quite straightforward: (1) because it is unjust (or wrong) to repay evil with evil (i.e., to repay an injury done to oneself with another injury or injustice); and (2) because he has agreed to abide by whatever judgments the state should pronounce in a court of law. In light of the first claim we claim we can see the point of what I shall call the "causal" or "consequential" argument below: it is supposed to show that by trying to escape Socrates would indeed be injuring the state. Thus, if he tries to escape, he is acting unjustly because of the moral principle against repaying evil with evil. What I call the "contractarian" argument is designed to show that Socrates has indeed agreed to abide by whatever judgments the state should pronounce in a court of law. Thus, trying to escape would also be wrong because of the moral principle requiring us to keep our just agreements.

The "consequential" argument fails not because Socrates would not be injuring the state by trying to escape but because the principle that we cannot repay injustice with injustice is fatally ambiguous. The "contractarian" argument fails, conversely, because while the principle that we should not break our just agreements is sound, it is unreasonable to suppose that Socrates has agreed to what the laws say he has agreed to. Before I can make these claims good, we shall have to look at the text in some detail.

5. Early in the dialogue Socrates sets the scene for his subsequent arguments by getting Crito to agree to three presuppositions that seem harmless enough but which turn out to be critical to the development of the subsequent arguments:

- (a) We must never act unjustly.
- (b) It is unjust to repay evil with evil, or injustice with injustice. Therefore, we must never do so.
- (c) We must carry out our just agreements.²²

I want to begin with a remark about the first of these presuppositions. Early in the dialogue Socrates gets Crito to acknowledge that "we ought never to act unjustly."²³ In light of Socrates' virtual identification of the notion of acting unjustly with acting wrongly (from the moral point of view), it might appear that Crito's concession amounts to no more than agreeing that it is immoral to act immorally. Without digressing into any deep questions concerning Plato's moral philosophy here, I want to point out that this first premise represents far more than an empty tautology. It expresses what I shall call Socrates' commitment to justice or right action. It indicates not merely that Socrates recognizes the truism just mentioned but also that he intends to act always as a just and upright man, whatever the consequences for himself. This is a rather obvious point, but we should bear in mind that Socrates' subsequent arguments will not be very convincing to anyone who does not go along with the first presupposition, for whatever reasons.

My second preliminary point is that there is a critical ambiguity in the notion of repaying evil with evil (injustice with injustice), and I'm afraid that Socrates' arguments sometimes get muddled because of it. Consider the distinction between:

- (1) Hurting someone simply because he has hurt me, where there is no other reason, excuse or justification for hurting him. (I shall call this repaying evil for evil in the "primary" sense or for its own sake.)

and

- (2) Hurting someone because he is hurting me, where the reason I am hurting him is that this is the only way I can stop him from hurting me. (I shall call this returning evil for evil in the "secondary" sense or for the sake of self-defense.)

I am inclined to grant (though even this, I think, is controversial) that returning evil for evil in the primary sense is unjust or morally wrong. If we can show that an action is an instance of returning evil for evil in this sense, then if we accept presuppositions (a) and (b) above, we must not perform that action. Often, however, an action may involve returning evil for evil in what I have called the "secondary" sense. I am inclined to think that whether returning evil for evil in this sense is unjust or immoral depends entirely upon whether or not the person inflicting the initial evil upon me is himself "justified" in doing so. If he is -- as he might be for a variety of reasons, as I shall show -- then returning evil for this (justifiable) evil may be itself unjust or immoral. If the evil initially inflicted upon me is not justified, however, then I cannot see how it can be wrong for me to resist it. We sometimes have to resist evil with means that would otherwise be evil but which are justified in the circumstances. I shall deal in a moment with the question of whether this is actually returning evil for evil.

Now it may seem paradoxical to suppose that anyone could ever be justified in inflicting injustice or evil upon another. But this is not so, for two reasons. First, by "evil" here Socrates does not mean "that which it is wrong to inflict," which would generate the paradox but, rather, "that which is an evil or is hurtful" which avoids the paradox. Secondly, and more importantly, we shall see below that it is arguable that under certain circumstances it is justifiable to treat people unfairly -- to do them an injustice. This certainly is paradoxical, but there are compelling arguments that we shall examine below in favor of accepting the unjust results of fair procedures. Indeed, I shall eventually show that this is precisely Socrates' point in the Crito.

To return to the main thread, it is crucial to distinguish the primary and secondary senses of repaying evil for evil for the following reason. If my account of this distinction is roughly correct, then the fact that an action can be appropriately described as "returning evil for evil" does not directly insure that the action is unjust or immoral, as Socrates sometimes suggests. In particular, if we are dealing with a case of returning evil for evil in the secondary sense, then it is essential

to determine whether the evil initially inflicted is justified or not.

The issue in the Crito is whether or not Socrates will be acting justly or rightly in trying to escape. It will not do for Socrates to point out that he would be "hurting" the state in some sense by escaping and then to claim, on the basis of presupposition (b) above, that it would be unjust for him to try to escape. Let us suppose, for now, that the "consequential" argument (see below) is sound and that Socrates would be injuring the state by trying to escape. Would escape be immoral then, because of premise (b)? It would be so only if Socrates would be repaying evil for evil in the primary sense. But this is not the case. At most he is returning evil for evil in the secondary sense. But in that case he is acting unjustly only if the state is justified in injuring him in the first place, and the burden of proof is on the state to demonstrate this, since Socrates is ex hypothesi innocent and has been unjustly convicted.

Socrates is simply confused, therefore, if he is arguing that, in light of presupposition (b) and the consequential argument, he has shown that he would be acting unjustly by trying to escape. The confusion arises out of the ambiguity of the notion of "repaying evil with evil." In presupposition (b) this phrase is used in the primary sense; the "consequential" argument merely shows that Socrates would be repaying evil with evil in the secondary sense. It is up to the state to show that it is justified in harming Socrates (by executing him despite his innocence) and the "contractarian" argument is supposed to show this. Thus, if the "contractarian" argument is sound, trying to escape would be unjust not only on presupposition (c) above but on presupposition (b) as well.

The shift from the "consequential" to the "contractarian" argument occurs when Socrates considers the following reply to the claim that he would be injuring the state by trying to escape: "But the state has injured me by judging my case unjustly."²⁴ Significantly, the laws do not deny this. They claim, instead, that it is irrelevant, because Socrates has agreed to abide "by whatever decisions the state should pronounce." This reply, of course, cuts two ways. Since it implicitly grants that

Socrates' claim is correct, it shows, on the one hand, that we are dealing with an apparent case of repaying evil for evil in the "secondary" sense. If this is so, then, as I have already shown, the burden of proof is on the state to show that the evil it is inflicting is justified and hence that Socrates is not justified in resisting that evil. The laws try to discharge this burden of proof by showing that they are justified because of Socrates' tacit agreement to abide by their final judicial judgments.

The "contractarian" argument is based on presupposition (c) above: we have a moral obligation to carry out our just agreements. The laws claim, quite simply, that Socrates has agreed "not in mere words, but in...actions" to abide "by whatever judgments the state should pronounce."²⁵ There are many questions we might raise with respect to this contention. I shall attempt to deal with the following: (1) Has Socrates in fact made some sort of implicit agreement with the state? (2) If so, is it the sort of agreement the laws describe and does it have the implications they say it has? (3) Would an "agreement" of this sort be a just and binding one? and finally, (4) Why would anyone ever make such an agreement, even tacitly?

Let us suppose for the sake of argument, what many thinkers would perhaps deny, that Socrates has -- at least tacitly -- made some sort of implicit "agreement" with the state. Is it the sort of agreement the Laws describe -- an agreement to "abide by whatever judgments the state should pronounce"? This seems terribly implausible, and thanks to Crito's intellectual limitations and his inclination to assent to just about anything his friend says, Socrates is not constrained to defend this claim explicitly in the dialogue. But I think we can imagine roughly how his defense of it would proceed if we ask ourselves the following question: why would anyone ever make such an agreement, and how could such a "blank-check" agreement, to abide by whatever the state decrees, be a "just" agreement?

The answer to the first question is suggested by Socrates' conception of the likely consequences of any other social arrangement. No answer to the second question is even suggested in the dialogue itself,

but an answer could be garnered from an examination of Plato's idea of justice in the Republic and elsewhere. We shall not have to turn to these sources, however, since the arguments of Rawls's that we shall consider below speak directly to this second question, the question of how it can be "just" or morally right to agree to perpetrate and endure injustice or what is itself morally wrong.

But what about the first question? Aside from the moral propriety of such an agreement, why does Socrates believe that it would be reasonable to agree to abide by whatever judgments the state should pronounce? Here to, Socrates' contentions are defended by Rawls, albeit obliquely and in an importantly qualified way. But first let us look at how Socrates might defend his implicit claim that a reasonable man would (and should) make an agreement like the one the laws describe.

Early in the dialogue Socrates -- in the mouth of the Laws of Athens -- sketches the first of his three major arguments against escape. This is what I called the "causal" or "consequential" argument above, an argument which depends upon the (allegedly) disastrous consequences that escaping would have. Thus, the Laws exclaim:

Tell us, Socrates, what have you in your mind to do? What do you mean by trying to escape, but to destroy us, the laws and the whole state, as far as you are able?²⁶

The force of the exclamation is apparently supposed to be explained when the Laws continue:

Do you think that a state can exist and not be overthrown, in which the decisions of law are of no force, and are disregarded and undermined by private individuals?²⁷

Socrates is close here to some form of extreme legalism, and for all that has been said so far it might be extreme legalism with respect to legislative decisions as well as extreme legalism with respect to judicial decisions. However, the following remark suggests that, explicitly at least, Socrates is only committed to the narrower form of legalism: "Much might be said," the Laws assert, "in defense of the law which makes judicial decisions supreme."²⁸

In order to construe the preceding remarks as a valid argument, we must interpret Socrates as making both a normative and a very strong empirical claim: (1) it is wrong to do anything likely to undermine the state, and (2) seeking to escape would be likely to undermine the state. Obviously, both claims need to be modified to be even tenuously plausible, but Socrates does not do so in the Crito. Before seeing how the political theory of John Rawls can be construed as an attempt to remedy these deficiencies, let me indicate where I think we have gotten thus far.

To make good on the preceding arguments, Socrates would have to show, first, why it would be wrong for him to do anything likely to undermine the state (including something like saving his own life when he has been unjustly condemned to death) and, secondly, why seeking to escape would in fact be likely to undermine the state. One obvious way of defending the normative claim would be the "utilitarian" way: arguing that more social benefits are achieved with a state than without one, and hence that undermining the state is wrong. Socrates does not use this argument, however, and relies instead on two other arguments: the "contractarian" argument (Socrates has agreed to abide by whatever decisions the state should pronounce) and the "relational" argument (Socrates owes it to the state to abide by its decisions and to avoid doing anything that would hurt it because the state has "reared" him -- he stands in a "child-parent," indeed a "slave-master," relationship to the state).

With respect to the first of these arguments we are really no further than when we started: supposing Socrates has made some sort of "agreement" with the state, why should we concede that it is such an extraordinarily rigorous one? It is at this point that the "consequential" argument is relevant once again. If the consequences of disobeying the state's laws, on the basis of individual private judgment, say, are in fact as bad as that argument suggests, then perhaps it would be reasonable to agree to abide absolutely by the state's (judicial) judgments, whatever they are (supposing all along, of course, that the state itself has remained more or less reasonably just).

6. The situation, then, is this. Two of Socrates' major arguments for

not escaping must be conjoined if either of them is to have any force: the consequential argument makes an empirical claim (about the likely consequences of disobedience to the law) that needs a normative claim to make it relevant (a claim about not doing what is likely to undermine the state); the contractarian argument provides us with the relevant normative claim (we shouldn't undermine the state because we have agreed not to) but is in difficulty over an empirical or at least pragmatic claim (the claim that in fact we have tacitly made such an agreement or that we would make it, given our ends and needs, if we had the chance). The theoretical resolution to this problem is obvious: if the consequential argument is correct in its empirical claim about the consequences of disobedience, then it seems that the contractarian argument can supply the relevant normative premise. For if the consequences of ever following one's private lights over social (or at least judicial) decisions are as likely to lead to disaster as the consequential argument suggests, then it would perhaps be reasonable to agree to forego individual private judgment for the greater good of a social decision procedure. Unfortunately, as obvious as this theoretical resolution is, as it stands it suffers from two insurmountable weaknesses.

To begin with, the empirical claim about the likely consequences of disobedience to the law is simply too strong. It is extremely implausible to claim, as Socrates must claim, that the slightest deviance from social (or at least judicial) decisions, even for the best of reasons, courts disaster. To be sure, there is an important and not implausible empirical claim in the offing here -- a claim about the dangers of individual private judgment and its likely consequences. But as we shall see below in our discussion of Rawls's defense of a similar argument, it is a carefully modified empirical claim.

This is just one of Socrates' problems, however. For suppose he were to grant that the consequences of disobedience are not always likely to be disastrous and that the maintenance of a healthy commonwealth is consistent with (indeed, perhaps dependent upon) a certain amount of conscientious, illegal dissent. If he made this concession, Socrates would

then no longer be able to retain the normative claim that he needs for his argument as a whole to go through. For the agreement to abide by whatever (judicial) decisions the state should make was reasonable only insofar as this agreement was necessary for the preservation of the state. If we grant, as it seems we must, that in fact the state can survive a certain amount of illegal dissent, then the most plausible "agreement" would be to abide by the state's decisions up to the point necessary for its preservation.

In short, Socrates faces a paradox: his normative claim, that we should do whatever is necessary to uphold a more or less just state, is quite plausible on a certain reading. But it does not follow that we have an absolute or unqualified obligation to abide by the state's (judicial) decisions, since it is not likely that absolute or unqualified obedience is necessary for the preservation of the state. In this case, Socrates seems to be making too strong an empirical claim. On the other hand, Socrates can weaken his empirical claim but try to make his argument good by strengthening his normative claim: we are required to abide by the state's decisions, whatever they are, whether compliance is necessary for the state's preservation or not. But this is, without further argument, too strong a normative claim. Thus, however we read him, Socrates is making either too strong an empirical or too strong a normative claim, and he cannot qualify the one without modifying the other -- with unhappy results for his argument.

Before turning to a treatment of these problems that is remarkable, among other reasons, for the way in which it picks up the thread of our discussion almost exactly where Socrates leaves off, I should like to anticipate one criticism that might be made of my treatment of his arguments in the Crito. Very briefly, it may be said that I've taken the letter of the text too seriously and the spirit not seriously enough. Thus, it might be conceded that the structure of Socrates' argument in the Crito is much the way I have described it, but that I have missed the real point of the dialogue. That point, it could be argued, is simply this: Socrates chose to play the Athenian "game," had his day in court, was unfortunately convicted, and yet is enough of a gentleman and fair-

minded citizen to accept the consequences. This, while in my opinion a rather farfetched interpretation of Socrates' arguments in the Crito, is in fact an argument for his accepting the consequences that we shall have to consider and that we shall consider when we evaluate Professor Rawls's arguments below. Here I should simply like to point out that however gentlemanly it might be for Socrates to play the sacrificial lamb, it is odd to think that he has a moral duty to do so or that he would be acting unjustly (or immorally) if he did not do so. And this is what any argument that would vindicate Socrates' position in the Crito must show. As a first step toward finding such an argument, I should like to examine Professor John Rawls's views about the nature and extent of our obligation to obey the law.

CHAPTER IV

RAWLS'S ARGUMENTS FOR OBEYING JUST AND UNJUST LAWS

1. For analytic purposes I shall interpret Rawls as making two claims in his later work: (i) citizens in a reasonably just constitutional democracy have a moral duty, other things equal, to comply with duly enacted laws and policies (this view is often expressed by saying that citizens in such a society have a prima facie moral duty to comply with duly enacted laws and policies); and (ii) citizens in a reasonably just constitutional democracy sometimes (and perhaps often) have a moral duty, on balance or all things considered, to comply with unjust laws and policies as well as just ones (this view is often expressed by saying that in such a society the injustice of a law or policy is not always sufficient to override one's prima facie duty to comply with duly enacted legislation).¹ Of course, Rawls is also making a third claim that we want to evaluate: that under certain circumstances, even when one is morally justified in disobeying the law in a reasonably just democratic society, one is nonetheless morally bound to accept the legal consequences for doing so. But since the best argument for this claim is an argument analogous to the argument for the first two claims, I want to postpone the analysis and evaluation of the former until we have gotten clear (in the present chapter) about the most plausible argument for (i) and (ii).

It seems that if (i) is correct, questions of obedience and disobedience to the law -- at least in reasonably just democratic societies -- are always moral questions. This evidently means that in such societies the fact that a certain course of conduct is proscribed or prohibited by a valid law is always a relevant moral consideration in deciding whether one ought to pursue that course of conduct or not. Needless to say, this fact is not always an overriding consideration -- this is what

is meant by calling the relevant duty a *prima facie* one or a duty "other things equal." Nonetheless, it appears to follow from the view we're concerned with that illegality always counts as one reason for not doing (or for doing, as the case may be) something that is prohibited or proscribed by law.

This may not seem terribly controversial at first blush, but it does appear to be a rather odd claim when we consider both how it is that something acquires the force of law in democratic societies and what an extraordinary variety of activities (or prohibitions on activities) have that force in such societies. To begin with the second point, we should have to note that smoking in most classrooms on the Ohio State campus, for example, is illegal. Hence, on the view in question smoking in those classrooms is at least prima facie immoral and in the absence of other (overriding) considerations is immoral on balance or all things considered.

Of course, this is a trivial example, but that's just its force. For the view in question has it that acting illegally is always prima facie immoral: that we are morally bound, other things equal, to comply with duly enacted laws and policies in societies like our own. Perhaps this is not intended to apply to laws like those against smoking in the classroom; but then we need to know how to determine when law has moral force and when it does not. Or perhaps it's true that the view in question applies in cases of this sort, but it's supposed to be obvious that other things are not equal in such cases. But then we need to know a great deal more about the phrase "other things equal" that appears in the first of the two claims that interest us.

A much larger issue is raised by the fact that the law is made, in societies like our own, with decision-procedures that are both majoritarian and representational (rather than direct). For present purposes I wish to emphasize only the first feature: the principle of majority rule. As we shall see, Rawls's argument is designed to show that, other things equal, we are morally bound to abide by the (valid) legislative decisions of a majority (or their representatives), even when we are in

the minority and even when we think the majority is acting unwisely and/or unjustly. Thus, the principle of majority rule is viewed not only as an obvious practical expedient but as a decision-procedure with considerable moral force behind it. I am not arguing against this feature of Rawls's view at this point, but simply trying to be sure we understand it and its implications. Perhaps some examples would be helpful.

A majority of the citizens in this country have seen fit, through their duly elected representatives, to prohibit the possession and use of marijuana, not to mention its sale. It is also illegal, in most states, for consenting adults to engage in homosexual acts of various kinds. It is even illegal in most states to engage privately in a variety of heterosexual activities, such as fellatio, cunnilingus, and sodomy. This is due of course to the moral views of the majority. Do those views have prima facie moral force for those of us who disagree with them and who wish to engage in some number of the illegal activities just mentioned? That is, is the fact that fellatio is illegal something I ought to be taking into account in my private life, if I wish to remain a righteous man? That it is allegedly unnatural is a fact that I am willing to take into account in the moral reckoning (though I reckon it not unnatural). But the fact that it is illegal? How can this be relevant to my private life? And we must remember that this is just one of an extraordinary number of valid laws that it seems odd to think of as relevant in my moral reckoning. What's more, many of them are laws that most of us either don't even know about or the details of which we have never bothered to consider. Is this moral irresponsibility on our part or just good sense, as I should have thought?

2. I should like to determine whether Rawls can get round difficulties of this sort, and what his argument for the two principal claims in question actually is: the claims that in a reasonably just constitutional democracy we have a prima facie moral duty to comply with duly enacted laws and policies and that in such a society it sometimes happens that we have a moral duty, on balance or all things considered, to comply with unjust as well as just laws and policies.² In order to understand Rawls's

strategy in his attempt to prove these claims, we must first note a distinction he uses in his discussion, a distinction that I am not sure I fully understand. This is the distinction between what Rawls calls "ideal" or "strict-compliance" theory and "nonideal" or "partial-compliance" theory. Roughly, the idea seems to be that in ideal theory we assume that everyone will always act in accordance with their moral and legal commitments, while in nonideal theory we try to take into account the fact that in real life everyone will not always act as they ought. Rawls believes that it is useful to do political theory from the point of view of ideal theory, in the first instance, and to make the appropriate adjustments for nonideal theory when we have gotten clear about the former (the situations of strict or perfect compliance).

Rawls's strategy is this. First he develops an argument concerning the moral basis of civil authority under the conditions (and subject to the assumptions) of ideal or strict-compliance theory. This is an argument that in a perfectly just democratic society we would be morally bound, other things equal, to comply with duly enacted laws and policies sans phrase. This argument does not bear on the second claim above -- since the problem of unjust laws and policies does not arise in ideal theory -- but it does suggest a rather straightforward defense of the first claim. Rawls then develops an argument from the point of view of nonideal or partial-compliance theory. This is an argument that even in an imperfectly but nearly just democratic society we would be morally bound to comply with duly enacted laws and policies even (up to a point) if they were unjust. This argument is supposed to prove both of Rawls's main claims as I am interpreting them. I shall begin by outlining and briefly evaluating the argument from ideal theory. The reasons for proceeding in this way will become obvious as we proceed.

The argument from ideal theory begins with a moral principle that Rawls calls "the natural duty of justice".³ This principle is formulated somewhat differently by Rawls in different places, and these differences turn out to be of some significance. Here, however, I wish to keep things fairly simple, so I shall say that according to Rawls the natural duty of justice requires us to support and uphold just institutions that exist

and apply to us and to further such institutions when they do not exist and when we can do so at no great cost to ourselves. For present purposes I am interested only in the first part of this principle, the claim that we have a natural duty to support just institutions that exist and apply to us.

Given the so-called natural duty of justice, Rawls goes on to argue that a perfectly just constitutional democracy would be a just institution in the relevant sense. It follows that citizens in a perfectly just constitutional democracy would have a moral duty to support and uphold it. But how do they become morally bound to duly enacted laws and policies in such a society?

I think that we must interpret Rawls as making an important empirical claim at this point: namely, that it is impossible, as a matter of fact, to maintain a perfectly just constitutional democracy without general compliance with some principle of institutional settlement -- the principle of majority rule, for example. If this is true, then since we have a moral duty to uphold such societies, it seems to follow that, for the most part, we have a moral duty to comply with its duly enacted laws and policies. I wish to summarize all of this as follows:

- (1) We have a natural moral duty to support and uphold just institutions that exist and apply to us -- institutions, that is, that satisfy the two principles of justice.
- (2) A just constitutional democracy is a just institution in the relevant sense -- that is, it is a social system that satisfies the two principles of justice.
- (3) Therefore, we have a moral duty to support and uphold a just constitutional democracy when it exists and applies to us.
- (4) Supporting and upholding a just constitutional democracy requires general compliance with some principle of institutional settlement for resolving legislative conflict -- the principle of majority rule, for example.
- (5) Therefore, we have a moral duty to comply with

the principle of majority rule in a just constitutional democracy that exists and applies to us. (In other words, we have a moral duty to comply with the duly enacted laws and policies of a just constitutional democracy that exists and applies to us.)

It is important to notice at the outset that the duty to support just institutions -- which for all we said above might very well be an important ethical principle in the moral philosophy of a utilitarian like Mill or an intuitionist like Ross -- is interpreted by Rawls as a duty to support and further arrangements that satisfy his own well-known principles of justice.⁴ That is, an ideally just institution simply is, for Rawls, an institution that satisfies these two principles. I have made this point explicit in the formal version of the argument from ideal theory above. Obviously, the so-called duty of justice becomes much more controversial as it is made more specific in this way. After all, it might be said, who wouldn't agree that we have a duty to support just institutions in some vague sense?

3. It would be impossible to consider here each of the many profound problems that the argument above faces.⁵ What I propose to do instead is show, first, that even if we concede premises (1), (2) and (4) above, Rawls still cannot validly infer (5), the conclusion he wants to establish. But I think I can suggest a move that will enable Rawls to remedy this difficulty. More interesting, as we shall see, is the fact that even the revised argument faces two further difficulties: it is so excessively abstract as to leave us wholly in the dark about the question of the nature and extent of our obligation to obey the law in situations of less than perfect justice; and it doesn't even mention the question of the moral force of unjust laws in imperfectly just societies. Of course, the argument from nonideal theory is intended to meet the latter difficulties. But first a word about the logic of the argument above, the argument from ideal theory.

Suppose we concede the substantive moral claim embodied in premiss (1), the stipulative claim made in premiss (2), and the empirical claim

suggested by premiss (4). It seems that Rawls still cannot validly infer (5), at least as it stands. The problem, of course, is that the notion of general compliance that (4) trades on is not adequately reflected in (5), the conclusion, and hence the argument is not even clearly valid. What's more, it is not clear that this difficulty can be easily remedied. The main idea, of course, is to get the notion of general compliance into (5), the conclusion.⁶ But notice that we cannot do this in any simple-minded fashion. We don't want to say, after all, that what the argument proves is that we have a natural moral duty generally to comply with duly enacted laws and policies in a perfectly just constitutional democracy. Aside from the unhappy fact that this is at best rather awkward English, it is surely less than what Rawls wants to say. We want, at the very least, regular (indeed, rather rigorous) compliance, and not just "general compliance" in some vague sense.

One obvious move is to reword (5) so that it states that in a perfectly just constitutional democracy we have what is usually called a prima facie moral duty (a moral duty other things equal, as Rawls puts it) to abide by duly enacted laws and policies.⁷ But the problem of the argument's formal validity remains: how exactly do we get from the empirical necessity for general compliance to the moral requirement of a prima facie duty to obey the law?

The intuitive idea, I take it, is that even from the standpoint of ideal theory it is necessary to bear in mind that without regular compliance our enterprise is doomed to failure. Not that anyone will be tempted to disobey the law for sinister reasons in ideal theory, or because of alleged injustice in the society's laws or policies. This is ruled out by the assumptions of strict-compliance theory. Still, Rawls says, even just men acting strictly on their (collective) sense of justice, will sometimes disagree as to what's the best policy under certain circumstances. They do not all have possession of the same set of facts, for example, and in any case different people will assess the relevance of various facts differently. For this reason, apparently, they would agree that justice will be more effectively served in the long run by com-

mitting themselves to an institutionalized decision-procedure for resolving their policy conflicts than by leaving what is to be done up to each individual acting by his private lights.

What we have, then, is a situation where men of good will agree that a commitment to some principle of institutional settlement (the principle of majority rule, for example) is requisite to keep their society going. And since they presumably view keeping it going as one of their most important natural duties, they will agree that it is, in general, more important to abide by the principle of institutional settlement than to have their own way in every case, as with the principle of private lights (roughly: "In the final analysis, always act by your own private lights.").

Now it may seem that with these intuitive remarks in mind we can solve the problem of formal validity mentioned above. That problem, it will be recalled, was this: on the one hand it seems quite difficult to alter (5) so that it will follow validly from (3) and (4); on the other hand the intuitive idea in the move to (5) seems fairly obvious. That is, what (4) is essentially saying is that unless people abide by the principle of institutional settlement for the most part, the just society cannot be maintained. Thus, what (5) should be saying is that for the most part we are morally bound to the results of the principle of institutional settlement, whether we happen to like them or not, due to our duty to support and uphold just institutions. Since we are not absolutely bound to these results, it is tempting to say that what the argument proves is that we have a prima facie duty, or a duty other things equal as Rawls puts it, to abide by duly enacted laws and policies in a just constitutional democracy. This all sounds very nice, but Rawls is still not out of the woods, as can easily be shown. The main difficulty is that the duty of justice is insufficient, in itself, to generate even a prima facie moral duty to obey the law.

Suppose that it is reasonable for some member of a given society to conclude that most of the other members will comply with the principle of majority rule and that noncompliance in his case will not

weaken the stability of the society in question. It seems to follow that he will not be violating the so-called duty of justice if he fails to comply with the principle of majority rule, for we are supposing that the society will be maintained whether he complies with it or not. What binds him to the principle of majority rule, then, if not the natural duty of justice?

I think that Rawls needs to introduce the principle of fairness, or the natural duty of fair play as he has also called it, in order to meet this difficulty.⁸ Any one citizen can fail to comply without endangering the society as a whole. But if enough citizens do so, then the society is in danger. What is true of one citizen, however, is true for any other citizen, as long as his fellows are complying. But now we face a situation which is familiar enough in cooperative ventures. The benefits of the venture -- indeed, its very viability -- are ensured only as long as most of the members comply most of the time. There is room for a certain amount of "free riding" as it is sometimes called. But how are we to determine who gets to ride free, since obviously everyone cannot? The principle of fairness suggests that no one is to ride free simply because he can get away with doing so. This would be unfair to the other members of the venture, who could just as well claim a right to ride free if it appeared that their doing so would not endanger the system as a whole. But this means that free-riding is forbidden not so much by the duty of justice, strictly speaking, but by the principle of fairness. Rawls's argument in A Theory of Justice does not bring this out clearly enough. We cannot generate even a prima facie moral duty to comply with duly enacted laws and policies without introducing the principle of fairness as well as the natural duty of justice, which was our starting-point.

If this is true, then we must insert another premiss into the preceding argument in order to generate the conclusion Rawls wants. I shall assume that the relevant moral principle is captured by premiss (5') in the argument immediately below -- call it the revised argument from ideal theory -- and that with it Rawls can make good on his claim that citizens

in a perfectly just constitutional democracy would have a prima facie moral duty to comply with all duly enacted laws and policies. The final argument from the point of view of ideal theory, then, is this:

- (1') We have a natural moral duty to support and uphold just institutions that exist and apply to us -- institutions, that is, that satisfy the two principles of justice.
- (2') A just constitutional democracy is a just institution in the relevant sense -- that is, it is a social system that satisfies the two principles of justice.
- (3') Therefore, we have a moral duty to support and uphold a just constitutional democracy when it exists and applies to us.
- (4') Supporting and upholding a just constitutional democracy requires general compliance with some principle of institutional settlement -- the principle of majority rule, for example.
- (5') If a social scheme requires general compliance with some principle of institutional settlement in order to be maintained, and if citizens are morally bound to uphold that scheme, then every citizen is morally bound to comply with duly enacted laws and policies, other things being equal.
- (6') Therefore, every citizen in a just constitutional democracy has a moral duty, other things equal, to abide by duly enacted laws and policies. (Alternatively: every citizen in a just constitutional democracy has a prima facie moral duty to abide by duly enacted laws and policies.)

Let us suppose that this argument is sound. How does it bear on the problem of the nature and extent of legitimate civil authority under less than perfect conditions? That is, under conditions other than those of ideal or strict-compliance theory?

4. What Rawls calls the argument from "nonideal" or "partial-compliance" theory is supposed to help us here. Rawls does not sketch this argument even roughly, but I take it that when worked out the argument from nonideal theory is supposed to do two somewhat different things: (i) prove that even in an imperfectly just society we have a prima facie

moral duty to comply with duly enacted laws and policies; and (ii) prove that this prima facie duty is heavily weighted, so that sometimes we are morally bound (all things considered) to comply with what we think, and think correctly, is an unjust law or policy. Rawls is inclined to say (or to imply, at any rate) that (i) goes without saying, in light of the argument from ideal theory, and that it's really (ii) that needs proving.⁹ But this is a mistake, for reasons that are worth explaining.

Rawls writes that "there is quite clearly no difficulty in explaining why we are to comply with just laws enacted under a just constitution".¹⁰ But this is an ambiguous remark, as I shall show, and it is possible that Rawls is making things easier on himself than he is entitled to, by exploiting that ambiguity. If he is merely saying that in a perfectly just society there would obviously be no difficulty in explaining why we are morally bound to comply with just laws enacted under a just constitution, then perhaps he's right and right for reasons of the sort suggested by the argument from ideal theory. However, it sometimes seems that Rawls means to say more than this when he says that there's no difficulty in explaining why we are morally bound to comply with just laws and policies enacted under a just constitution. For his remarks sometimes suggest the idea that in nonideal theory, when we are dealing with imperfectly just institutions, we can separate just laws and policies from unjust ones, give a quick proof of why we are morally bound to the former (by alluding to the argument from ideal theory), and then formulate a somewhat lengthier account of how we come to be bound to (some of) the latter as well. But this won't do, for a number of reasons.

To begin with, all that the argument from ideal theory proves, if it proves anything, is that citizens in a perfectly just society would have a moral duty (other things equal) to obey just laws and policies enacted under a just constitution. When we turn to imperfectly just societies, like our own, it will not do for Rawls to single out the just laws and policies, refer us back to the argument from ideal theory, and contend that that argument explains why we are morally bound to at least the just laws and policies in the imperfectly just society. He cannot do this because all the earlier argument established was that we would be

morally bound to just laws and policies in a perfectly just society. It is by no means obvious that this same account holds for just laws and policies in an imperfectly just society. It may be that in imperfectly just societies we have no moral duty to obey the law at all.

But Rawls's claim that there's no problem in explaining how we become morally bound to just laws and policies, even in imperfectly just societies, is misleading for another reason as well. For it suggests that in a given society we can distinguish just and unjust laws, for example, and then show how we are obviously bound to the former and at least arguably bound (up to a point) even to the latter. But as we shall see, the thrust of Rawls's main argument for complying with at least some unjust laws and policies is that we cannot determine, with respect to a large and important number of cases, whether or not a given law or policy is "really" unjust. That is, this is a distinction that it is dangerous and, in any case, often impossible to make with respect to a given law or policy, according to Rawls. But then Rawls himself is hardly in a position to rely on this distinction in order to pass off his claim about obeying just laws and policies as though it were a truism or followed in some obvious way from his remarks about ideal or perfectly just societies.

In short, the argument from nonideal or partial-compliance theory has got to be an argument about laws and policies (in a reasonably just society) generally; it cannot be a two-part argument, where one part proves that we are morally bound to just laws and policies in such a society and the other proves that we are even bound (up to a point) to unjust laws and policies. I believe that an argument of the appropriate sort can be found in Rawls's book, but in large part it is an argument that I have put together from remarks he makes in different places. As I see it, the central argument in nonideal theory is parallel, up to a point, to the argument from ideal theory. But it becomes quite another argument at a certain point, as I shall show.

If the argument from ideal theory is sound, then in a perfectly just constitutional democracy we would be morally bound to comply with

duly enacted laws and policies in such a society -- laws and policies which will, by assumption, always be just. The difficulty that the argument from nonideal theory has to meet is simply this: it is extremely unlikely that we shall ever find ourselves in a perfectly just society, so what's the relevance of the so-called natural duty of justice? At least two moves are open to Rawls at this point.

On the one hand, he might simply remind us that there's more to the natural duty of justice than has thus far been acknowledged. In addition to supporting just institutions that exist and apply to us, we are morally bound to foster the development of such institutions when they do not exist, at least if we can do so at no great cost to ourselves. Hence, we might expect Rawls to argue that we are morally bound to support imperfectly just societies when doing so is necessary for the development of perfectly just societies. And if supporting imperfectly just societies requires obeying duly enacted laws and policies in such societies, then Rawls might argue that we are morally bound (other things equal) to comply with duly enacted laws and policies even in imperfectly just societies and even when it's certain, for all practical purposes, that not all of those laws and policies will be just.

Rawls does not follow the strategy suggested by the preceding remarks, however, and he does not explain why. Perhaps he feels that it's manifestly implausible to claim that supporting imperfectly just societies is required for fostering perfectly just societies. Or perhaps he recognizes that, since the duty of justice requires us to foster the development of perfectly just societies only if we can do so at no great cost to ourselves, it is not likely that we are committed to doing this by complying with the laws and policies of an imperfectly just society. After all, this is likely to involve considerable "cost" to ourselves, both in terms of personal inconvenience and in terms of finding ourselves forced to obey laws and policies that we consider palpably unjust. It is important to remember that Rawls could have advanced some argument like that suggested above, however, since the argument that he does advance is not entirely unrelated to it and certainly makes more sense when we think

of Rawls's somewhat abstract but nonetheless very serious commitment to the development of perfectly just societies -- societies, that is, that realize the principles of justice perfectly.

The argument that Rawls in fact uses to make good on his central claims in nonideal theory is related to the argument from ideal theory in an obvious way. Since it's unreasonable to expect to find perfectly just institutions in the real world, it is plausible to suppose that we nonetheless have a natural moral duty to support institutions that come reasonably close to meeting the standards set by the two principles of justice. Thus, the argument from nonideal theory involves a significant (but not very surprising) amendment to the natural duty of justice as we have been thinking of it so far; we might put this principle as follows:

We have a natural moral duty to support and uphold reasonably just institutions that exist and apply to us -- institutions, that is, that come reasonably close to satisfying the two principles of justice.

With this principle as a starting-point, it is easy enough to see how the argument from nonideal theory is to proceed. Rawls need only claim that a reasonably just constitutional democracy would be a reasonably just institution in the relevant sense, and it follows that we would be morally bound to uphold a reasonably just constitutional democracy that existed and applied to us. Given an empirical claim analogous to that made in premiss (4') of the argument from ideal theory, as well as some principle of fair play like that embodied in premiss (5') of the earlier argument, it appears that we can deduce at least one of the claims Rawls wishes to make: namely, that even in imperfectly just democratic societies there is a moral presumption in favor of complying with duly enacted laws and policies. I wish to summarize this argument formally as follows:

- (1") We have a natural moral duty to support and uphold reasonably just institutions that exist and apply to us -- institutions, that is, that come reasonably close to satisfying the two principles of justice.
- (2") A reasonably just constitutional democracy is

a reasonably just institution in the relevant sense -- that is, it is a social system that comes reasonably close to satisfying the two principles of justice.

- (3") Therefore, we have a moral duty to support and uphold a reasonably just constitutional democracy when it exists and applies to us.
- (4") Supporting and upholding a reasonably just constitutional democracy requires general compliance with some principle of institutional settlement for resolving legislative conflict -- the principle of majority rule, for example.
- (5") If a social scheme requires general compliance with some principle of institutional settlement in order to be maintained, and if citizens are morally bound to uphold that scheme, then every citizen is morally bound to abide by duly enacted laws and policies, other things being equal.
- (6") Therefore, every citizen in a reasonably just constitutional democracy has a moral duty, other things being equal, to abide by duly enacted laws and policies. (Alternatively: every citizen in a reasonably just constitutional democracy has a prima facie moral duty to abide by duly enacted laws and policies.)

5. Now this argument faces difficulties of at least four kinds. To begin with, it faces all the difficulties that the argument from ideal theory faced. Thus, it would need to be shown that we do indeed have a natural moral duty to uphold reasonably just institutions and that such institutions are those that come reasonably close to satisfying Rawls's principles of justice. Similarly, it would have to be shown that it is in fact impossible to maintain a reasonably just constitutional democracy without relying on some principle of institutional settlement such as the principle of majority rule. But I shall not pursue these problems here.

There is a somewhat different problem connected with premiss (1") in the argument from nonideal theory, however. This is the problem of articulating just what it means to say that we have a moral duty to support reasonably just institutions as well as perfectly just ones. Even if we suppose that we know what it would be like for a society to satisfy the two principles of justice perfectly, it's not entirely clear what

someone would have in mind in speaking of a society that came "reasonably close" to doing so.

Rawls's idea, I take it, is roughly this. A social arrangement comes reasonably close to satisfying the two principles of justice if it comes as close to realizing those principles as it is reasonable to expect under the circumstances. Thus, if a given society tends to maximize political freedom and equality of opportunity, and tends to distribute social and economic inequalities in a way calculated to improve the lot of those least fortunate in that society, then, even if it does not do so perfectly, we have a natural duty to uphold that society and to do what is required of us to keep it going. Rawls is assuming as well, of course, that the society in question is coming fairly close to maximizing these things, at least as close as any alternative and, again, as close as it is reasonable to expect under the circumstances. In other words, while it may well be conceivable that the society in question could be doing better, it would be unreasonable to expect it to do so under the circumstances (the circumstances of normal political life, that is). At the very least, this must mean that that particular society is likelier than any available alternative social arrangement to realize the principles of justice, even imperfectly. On this interpretation, it seems to me, premiss (1") is not entirely implausible.

But of course there's a third difficulty connected with the argument from nonideal theory, a difficulty which is independent of the one just mentioned. For how can we say, in an actual case, when a given society is as just as it's reasonable to expect, under the circumstances? That is, even if we suppose that Rawls's principles of justice are our standards for assessing the relative justice of a social arrangement, and even if we agree that a society would be "reasonably just" if it came reasonably close to satisfying these principles, it turns out to be very difficult (if not impossible) to get any sort of agreement, in practice, on the question of whether or not a given society is indeed coming as close to satisfying the principles of justice as it is reasonable to expect. Is our own society, for example, a reasonably just one in Rawls's

sense, so that all U.S. citizens are morally bound to comply with its duly enacted laws and policies? It's not just that it's difficult to get general agreement on the correct answer to this question. The social and political difficulties of the past decade suggest that it may be impossible.

Thus, even if Rawls has proven that there would be a moral presumption in favor of obedience to the law in a reasonably just constitutional democracy as he conceives it, this is not a very helpful result. It will probably always be a matter of serious controversy as to whether or not a given society is as just as it is reasonable to expect, even if we accept Rawls's theory of justice. Hence, it will apparently always be a matter of genuine controversy as to whether or not we are in general morally bound to abide by duly enacted laws and policies in a given society, even if we accept Rawls's "theory of right" or of the moral basis of civil authority.

But the largest difficulty with the argument thus far -- and the one I wish to follow up in the remainder of this chapter -- is this: in its present form the argument does not speak to the problem of unjust laws and policies and of the extent to which, theoretically at least, citizens in a reasonably just constitutional democracy would be morally bound to submit to them. Thus, we need another premiss to get the conclusion Rawls wants: that in a reasonably just constitutional democracy we have a prima facie moral duty to abide by unjust as well as just results of the principle of majority rule.

6. Rawls recognizes that in a reasonably just constitutional democracy it will sometimes happen that we will be confronted with a duly enacted law or policy that we think, and think correctly, is unjust. The problem in such cases is that we seem to be faced with a conflict of duties: on the one hand we are morally bound to comply with duly enacted laws and policies in such a society (at least if the argument above is sound); on the other hand, it is plausible to say that we have a natural duty (and in any case a right) to resist injustice. What do we do, then, in cases of duly enacted laws and policies that we believe, correctly let us

suppose, to be unjust?

The simple answer is to say that since the duty to comply with duly enacted legislation in a reasonably just society is only a prima facie duty or a duty "other things equal", we are never morally bound to comply with an unjust law. That is, we might say that the injustice of a given law or policy always overrides our moral duty (other things equal) to comply with duly enacted laws and policies, and hence that we are never morally bound to comply with unjust laws and policies, even in a more or less just constitutional democracy. My own view is that this simple answer is essentially correct. Rawls rejects this answer, however, and in this respect his views are characteristic of most recent liberal (not to mention conservative) political theorists. What reason is there for supposing that such thinkers are correct, and that we are sometimes morally bound to go along with unjust laws and policies in more or less just democratic societies?

It is tempting to think that Rawls can get the result he wants simply by altering premiss (5'') above in an obvious way (see (5*) below). On this view, which was Rawls's view in an early paper and is still favored by many political theorists, we are morally bound to unjust laws and policies for the same reason that we're morally bound to the results of the principle of majority rule generally.¹¹ Others have done their part, hence we must do our part when it comes our turn, even if this means complying with unjust laws. In other words, we are morally bound to duly enacted but unjust laws and policies because we are morally bound to duly enacted laws and policies generally, other things being equal, and hence we must obey putatively unjust laws and policies just as we must obey inconvenient ones, say, when it comes our turn. To do otherwise would be to act unfairly. Hence fairness (that is, justice) requires us to do what would otherwise be unjust or unfair.

We can capture the thrust of the preceding remarks simply by altering (5'') so that it reads as follows:

- (5*) If a social scheme requires general compliance with some principle of institutional settlement in order to be maintained, and if citizens

are morally bound to uphold that scheme, then every citizen is morally bound to abide by duly enacted laws and policies, other things being equal, even, up to a point, if they are unjust.

We might then infer (6*) or something like it:

- (6*) Therefore, every citizen in a reasonably just constitutional democracy has a moral duty, other things being equal, to abide by duly enacted laws and policies, even, up to a point, when they are unjust. (Or: every citizen in a reasonably just constitutional democracy has a prima facie moral duty to abide by duly enacted laws and policies, even when they are unjust.)

The trouble with construing the argument for obeying unjust laws in this way is not merely, nor even primarily, that it is ad hoc. Rather, the problem is that the argument from (5*) proceeds as though the injustice involved in "doing one's part" in certain cases -- cases where the law or policy is unjust -- is not a relevant difference with respect to the general notion of doing-one's-part, other things equal. That is, at first blush it seems that the fact that a law or policy is unjust implies that other things are not equal, and hence that the principle that I ought to do my part, other things equal, simply does not apply.

But this tack is mistaken for even more profound reasons. For when properly understood, it changes the very direction of Rawls's argument for the moral basis of civil authority. This is because premiss (5"), and the principle of fair play generally, plays only a subordinate role in the argument for complying with duly enacted laws and policies in a reasonably just constitutional democracy. The basic thrust of both the argument from ideal and the argument from nonideal theory is that we're morally bound to comply with duly enacted laws and policies because without general compliance the system cannot be maintained. And maintaining the system is one of our most serious moral duties. The idea in the earlier arguments was not "Others have done their part, therefore you must do your part when it comes your turn", but "People must obey for the most part, and you're no exception unless you can show why."

In other words, the principle of fairness does not account for why, in general, we are morally bound to obey the law in a more or less just democratic society. Rather, the natural duty of justice accounts for our duty to comply with duly enacted laws and policies, via the notion of the principle of majority rule and doing what is required of us to uphold just institutions. The principle of fairness was introduced only as a means of handling the problem of "free riders" and to make the move from (4') to (6') valid. To think of the principle of fairness as lying at the heart of the matter, as the argument for (5*) does, is to make Rawls a social-contract or "consent" theorist in a very real sense. Although this would be a correct description of the early Rawls, it is certainly not true of the later Rawls, as I have remarked above.¹²

7. It will be convenient, for analytic purposes, to note that the view we've been considering is subject to still another important criticism. Rawls remarks early in his book that "an injustice is tolerable only when it is necessary to avoid an even greater injustice."¹³ But this implies that an injustice would not be morally tolerable, much less required of us, merely because it happened to be our turn to obey the law and the law in question happened to be unjust. Quite the contrary. One would be justified in doing one's turn, if doing so involved one in perpetrating injustice, only if doing one's turn, and thus perpetrating the injustice, were required to prevent an even greater injustice. But what could that (latter) injustice be? The answer to this question brings us to the heart of Rawls's real argument for obeying unjust laws. Very roughly, it is an argument which depends upon the empirical claim that abiding by (some) unjust laws and policies is necessary if we are to uphold a reasonably just constitutional democracy, rather than on the formal claim that one must do one's part, in such a society, when it comes one's turn.

At first blush the premiss Rawls needs is an empirical one of the following sort: an otherwise just society simply cannot be maintained unless we agree to abide by some unjust laws and policies. In other words, the principle of majority rule will work, as a procedural device for

resolving legislative conflict, only if we abide for the most part by both its just and, up to a point, its unjust results. For the sake of simplicity, I shall interpret Rawls as adding the following premiss to his argument above (see section 4) in order to generate the appropriate conclusion:

- (7") Maintaining a reasonably just constitutional democracy requires general compliance not only with duly enacted laws and policies that happen to be just, but with duly enacted laws and policies generally -- even, up to a point, when they are unjust.

If Rawls can prove this claim, then he can infer (8"), the conclusion he wants:

- (8") Therefore, every citizen in a reasonably just constitutional democracy has a moral duty, other things equal, to abide by duly enacted laws and policies even when they are unjust.
(Or: every citizen in such a society has a prima facie moral duty to abide by duly enacted laws and policies, even when they are unjust.

Can Rawls prove premiss (7")?

Notice, to begin with, that the empirical claim embodied in (7") is surely not that a reasonably just society cannot be maintained unless its occasional unjust laws and policies are regularly complied with as such. Any society which could not survive if every one of its unjust laws and policies were violated -- but none of whose just laws and policies were violated -- would hardly be a clear (or even a plausible) instance of a just or reasonably just society. I take it that the thrust of Rawls's argument for complying with unjust laws is not that we are committed to unjust results as such, then, but that we are committed to acting in accordance with the principle of majority rule generally, even when its results are unjust.

But why? Why, when the results are palpably unjust, must we abide by the principle of majority rule? I believe that Rawls's thought is this: if we could actually know when this were the case (i.e., identify unjust laws publicly and agree collectively that they were unjust), then perhaps we would never be required to go along with them. But we have no Moral Oracle, as it were, of the sort that could conclusively resolve the

relevant questions, and hence the principle we would be acting on would be a principle of private judgment of some sort: for example, "Judge for yourself when a law is unjust and hence not to be obeyed." Why is Rawls opposed to such a principle?

Suppose we interpret Rawls as claiming that a reasonably just society simply cannot be maintained without rigorous compliance with some principle of institutional settlement, even when its results are unjust. Without such compliance, that is, a society would simply collapse. This is not unlike the argument advanced by Socrates in the Crito. It is a bad argument, however, and hence it is probably mistaken to interpret Rawls as making quite the same claim. But it is worth noting just what Rawls's claim would be on this interpretation, and why it would be mistaken.

Socrates argued in the Crito that a society simply cannot endure in which decisions of law have no force and are undermined by private individuals. There are at least two serious mistakes in this line of reasoning. To begin with, it's false to say that decisions of law have no force for an individual as long as he reserves the right to disregard them when he thinks they're unjust. Such decisions will certainly have force for him, morally speaking, when he believes they're just, even if they're terribly inconvenient or obviously not in his own interest. What's more, even putatively unjust laws, laws which in the final analysis he is perhaps unlikely to obey, might have more "force" for him than other sorts of institutionalized injustice. That is, the fact that apparent injustice is embodied in a law will give this person pause (make him think things out a little more carefully before acting by his private lights) in a way that the ordinary injustices of everyday life do not. He may feel no qualms about resisting injustices of the latter sort; but he may well think very carefully before concluding that a duly enacted law is unjust and hence to be resisted. In this sense even laws that a person is likely to disobey, in the final analysis, have a certain force for him qua laws.

But it is even more important to notice that Socrates' claim in the

Crito is simply too general. It may well be true that if every citizen in a given society always acts by his private lights, rather than by some just principle of institutional settlement, that society cannot long endure. This is not the kind of case that's relevant, however. We're concerned with a situation where citizens quite willingly acknowledge the right of the majority to legislate, and have their way, with respect to the countless practical issues that must be settled if a just and efficient society is to be maintained. The question is simply whether they must also surrender their right to assess questions of justice and injustice by their own lights, at least up to a point, if their society is to survive.

Socrates makes it too easy for himself in the Crito, in other words. We need not deny that complete lawlessness is likely to undermine the state in order to defend the claim that allowing a certain amount of individual discretion with respect to putatively unjust laws is consistent with keeping society going.

Of course, Rawls's critic must not be allowed to make it too easy on himself either. It may well be true, as a straightforward matter of empirical fact, that even if we allow private judgment to prevail only in cases of putative legislative injustice, we will end up with more private dissent from social policy than society can safely tolerate. This is an important empirical question which has not been adequately studied, so far as I know. We all have our own intuitions about such matters, but these are practically worthless. What we need is to encourage social scientists to do some hard research on the question of the likely effects of allowing (and even encouraging) private judgment on any issue that seems to one, upon reflection, to raise serious questions of legislative injustice.

It can be argued, however, that to cast the discussion in this way is to miss an obvious point. Whether or not the state can in some sense continue to exist side by side with a great deal of individual judgment over and against institutionalized social decision-procedures, it is surely obvious that a society in which private judgment regularly overrides public

policy, even if this is only on questions of justice and injustice, will be less efficient than a society in which people abide fairly rigorously by some principle of institutional settlement.

For present purposes I wish to concede this claim. The kind of society Rawls's critic is apparently advocating (see below) would presumably be less efficient in some straightforward sense than the alternative Rawls favors. But suppose we interpret Rawls as making this fairly obvious claim. That is, suppose we interpret him as claiming not that society cannot survive unless there is general compliance with some principle of institutional settlement, even on questions of justice and injustice, but merely that society will obviously be less efficient unless some principle of institutional settlement is generally agreed to have more moral force than one's own private moral judgments.

Rawls's empirical claim would then be far more plausible, but surely it is one which he would not wish to defend, at least in its present form. For Rawls is committed, above all, to defending justice against utility, to put it roughly. Suppose our society will be more efficient if people more or less regularly squash their scruples with respect to moral matters in favor of institutionalized decision-procedures. Isn't the inefficiency of the one system outweighed by the fact that in the alternative system we give wider scope to a person's balanced judgments of the right and the just?

The answer to this last question is obvious, I think, when we remember Rawls's maxim that an injustice is tolerable only when it is required to prevent an even greater injustice. On this view it is not the bare inefficiency of the principle of private judgment that requires us to favor the principle of majority rule. Rather, it is the possibility that following the principle of private judgment (with respect to putative injustice) will lead to greater injustice, in the long run, than following the principle of majority rule, even when its results are unjust. But this leads to a new interpretation of Rawls's thought in premiss (7").

8. According to the new interpretation Rawls is claiming that somehow

society will be less effectively just (in the long run) if we refuse to comply with unjust laws and policies than if we agree to do so, within limits. Aside from its other merits, this is the only interpretation of Rawls's empirical claim which is at all plausible, on the one hand, and is consistent with his dictum that an injustice is tolerable only when it is required to prevent an even greater injustice, on the other. We now need to ask what sense can be made of this new claim and what can be said in its behalf.

Obviously, the intuitive idea is that regular, nonselective compliance with duly enacted laws and policies is necessary to prevent some greater injustice than the injustice involved in complying with occasional unjust laws and policies. That is, on the new interpretation we are to comply with duly enacted laws and policies -- even, within limits, when they are unjust -- because doing so is necessary to prevent an even greater injustice than the injustice the law will sometimes sanction. But what injustice, exactly, is the latter injustice (the injustice of unjust laws and policies as well as the injustice of complying with them) supposed to help us avoid?

I think Rawls has at least two different sorts of things in mind here. On the one hand, certain of his remarks suggest that he believes that we simply cannot manage a democratic regime unless we renounce the principle of private rights for the most part and agree to abide by the principle of majority rule.¹⁴ Rawls is also thinking along lines such as this, I believe, when he remarks that our commitment to the principle of majority rule is necessary to gain the advantages of an effective legislative procedure.¹⁵ These sorts of considerations suggest that we would do best to interpret Rawls as arguing in (7") that it's impossible to manage a democratic regime (and secure the benefits that accrue to it) without general compliance with the principle of majority rule, even, within limits, when its results are unjust.

But it would be a mistake to take this interpretation too simply-mindedly. In particular, it would be mistaken to think of the injustice we avoid by giving up the principle of private judgment as some fairly

specific instance of massive injustice that we ward off by regular compliance with the principle of majority rule. (This was Socrates' mistake in the Crito, when he argued that if he refused to go along with the unjust results of a just decision-procedure he would be contributing to the very dissolution of an otherwise just regime.) Rather, I take it that Rawls's idea is roughly the following. Society will work more effectively in securing the basic liberties for every citizen -- and in seeing to it that inequalities are to the advantage of all (especially the least advantaged) -- only if the principle of majority rule is fairly rigorously complied with, even when its results are unjust in certain cases. That is, general compliance will tend to realize these ends more effectively than, say, selective compliance -- compliance that is fairly regular but that stops whenever complying would mean going along with what one believes to be an unjust law or policy.

But of course now we must ask a number of rather basic questions. To begin with, what evidence is there that this new empirical claim is true? And how, even in a general way, would one go about collecting evidence concerning such a claim, one way or the other?

Obviously, we cannot deal with these questions in any detail here, and in fact we really don't need to. Our main interest is in the argument for paying the penalty that can be constructed by analogy to Rawls's argument above. Hence we could just as well concede everything Rawls has said so far, since we have what we need: the bare bones of an argument for restricted legalism with respect to legislative decisions. We could simply get on with it, then, and ask where this reasoning leads when it comes to the judicial system and paying the penalty for justified civil disobedience.

But I cannot forbear from concluding this chapter with some tentative remarks about Rawls's argument for complying with unjust laws. My main complaint is that Rawls never tells us what sorts of laws he has in mind in this connection. Moreover, when we try to imagine concrete cases of laws that would apparently be within the bounds of what Rawls calls "tolerable injustice", it seems clear that in many such cases we

would in fact not be morally bound to go along with legislation of this sort.

9. Suppose we ask when people ought to comply with laws that they believe to be unjust, on Rawls's view, and why they should do so. We are in the curious position of knowing roughly why Rawls would have people comply with what they believe to be unjust laws but of not knowing when -- that is, of not knowing exactly what sorts of cases Rawls has in mind as instances of "tolerable injustice", as he puts it.

Of course, we know from Rawls's theory of justified civil disobedience roughly when a given law or policy is beyond the bounds of tolerable injustice, and hence the following idea suggests itself. Suppose we sketch the sorts of cases that qualify as clear cases of intolerable injustice for Rawls, and then try to determine what might be considered tolerable injustice by working backwards, by imagining cases that don't qualify. I shall attempt to do this below, but first a serious caveat: it is not at all clear that Rawls means to say that when neither civil disobedience nor conscientious refusal is justified in a reasonably just constitutional democracy, then noncompliance is out of the question. These are not the only forms of justifiable dissent on Rawls's view, perhaps even in a reasonably just democratic society. Still, one wants to know how Rawls feels about the moral force of certain putatively unjust laws that do not meet the requirements of either the theory of justified civil disobedience or conscientious refusal.

I shall contend that if Rawls maintains that we are morally bound to obey the law in such cases, then he is simply mistaken. Alternatively, if he would agree that noncompliance is justified in the sorts of cases that I shall describe, then I contend that his claim that sometimes we are morally bound to comply with unjust laws is uninteresting.

The general theory of noncompliance that Rawls develops in his book has two parts: the theory of civil disobedience and the theory of conscientious refusal. These theories are supposed to tell us roughly when noncompliance (with duly enacted laws and policies) would be morally

justified, even in a reasonably just constitutional democracy. In other words, they indicate roughly when the duty (or at any rate the right) to resist injustice overrides the duty to comply with duly enacted laws and policies, just and unjust alike.

Briefly, the theory of civil disobedience sets three conditions for justified civil disobedience as Rawls conceives it.¹⁶ (1) We must be dealing with alleged injustices that are "serious infringements of the first principle of justice" and/or "blatant violations of the second part of the second principle of justice". (Rawls explicitly excludes an appeal to the first part of the second principle of justice, the so-called "difference principle". I discuss this point below.) (2) Normal appeals must have been made through the standard modes of legal redress. (3) The dissenters must be careful to see to it that serious disorder is not likely to follow in view of the fact that other groups, similarly situated, have an equal right to protest.

Conscientious refusal is justified in two sorts of cases: (1) cases that violate a nation's jus ad bellum (that is, cases of unjust wars); and (2) cases that violate a nation's jus in bello (that is, cases of unjust warfare in an otherwise just war).¹⁷ The latter, according to Rawls, are cases where a man could say that "all things considered, his natural duty not to be made the agent of grave injustice and evil to another outweighs his duty to obey."¹⁸

It is obvious that Rawls believes that managing a reasonably just democratic society, and securing the advantages of such a society, requires compliance with a certain amount of unjust legislation. This, after all, is his central claim in section 53 of his book. But what sorts of cases does Rawls have in mind here? If we judge injustice in terms of the principles of justice, then Rawls has already conceded that we are not morally bound to go along with "serious" violations of the first principle of justice or "blatant" violations of the second. What sorts of unjust laws are we left with, then, as instances of legislation that is within the bounds of tolerable injustice?

It seems that there are at least three sorts of cases: (1) any law

that seems to violate the first part of the second principle of justice (assuming, of course, that the society as a whole is reasonably just and hence comes reasonably close to satisfying the difference principle); (2) any law that seems to violate the second part of the second principle of justice but is not a "blatant" violation of that principle; and (3) any law that seems to violate the first principle of justice, but which is not a sufficiently "serious" violation of that principle to justify noncompliance.

There is a great deal of legislation in our own society, I think, that violates only the first part of the second principle of justice, and which we are morally bound to go along with, according to Rawls, despite the fact that some of the most compelling reasons for disobedience to the law in our society have to do with this principle. I think a good case could be made against Rawls in cases of this sort, but I shall not try to do so here. Rather, I wish to concentrate on putative violations of the first principle of justice, according to which every citizen is entitled to maximum equal liberty consistent with a like liberty for all, and the second part of the second principle, which guarantees fair equality of opportunity. These are the most important sorts of cases, I believe, because it is precisely when the law seems to infringe upon my basic liberties that it seems to me to lose whatever moral force it may ordinarily have.

10. The sorts of cases I have in mind here are cases in which the law seems to interfere with a person's right to live as he or she sees fit, as long as in doing so he or she is not interfering with the rights or well-being of others. Thus, the majority of citizens in most states have seen fit, through their duly elected representatives, to prohibit the sale, possession and use of marijuana. It is also illegal, in most states, for consenting adults to engage in certain homosexual activities in private. It is even illegal, in many jurisdictions, to engage privately in certain heterosexual activities, such as sodomy, fellatio and cunnilingus. It is illegal in many states to procure a prostitute, and illegal for prostitutes to solicit customers. And so on for the whole list of so-

called "paternalistic" legislation and legislation that creates what are sometimes called "crimes without victims".

What is Rawls's position with respect to cases of this sort? Is it morally wrong to smoke marijuana, procure a prostitute, perform cunnilingus, etc., when (and because) these activities are illegal? If these things are not wrong on Rawls's theory, then what is the point of his claim about obeying unjust laws? If they are, how can his theory be correct?

I should like to begin by sketching very roughly my own position with respect to laws of this sort. When I believe that a law intrudes (unnecessarily) upon one or more of my basic liberties, I disregard it, and I am willing to allow others to do likewise (not merely with respect to the laws I find objectionable, of course, but with respect to laws that they feel intrude unnecessarily upon one or more of their basic liberties).

Consider the laws against possession and use of marijuana, for example, and any other illegal drugs usage of which has no demonstrable ill-effects on society. I disregard these laws and I do so because I feel they violate my right to privacy as well as my right to live my life as I see fit, as long as doing so is not harmful to others. Is there any sense in which my doing so endangers the stability of American society?

Obviously, what's worrying Rawls here is not so much the possible ill-effects of my own illegal activities, but what he considers the likely ill-effects of everyone's doing as I do. But of course he's not worried about the ill-effects of everyone's smoking marijuana, or even of everyone's breaking the law in similarly harmless cases. Rather, what he's worried about, I take it, are the likely effects of everyone's acting on a general maxim like my own: when one believes that a law intrudes unnecessarily upon one or more of his basic liberties, one may disregard it.

Suppose we distinguish two classes of laws that people might believe

fall under this principle and hence which they would be likely to disregard if they accepted it. The first class of cases will be the general class described above: laws which seem to interfere with a person's right to live as he or she sees fit, as long as in doing so he or she is not interfering with the rights or well-being of others. I have already mentioned a number of examples of laws of this sort. It would be preposterous, I believe, to contend that generalized disobedience with respect to laws of this sort would have bad effects on society's stability. The proof of this is that in fact most people who feel that these laws interfere with one or more of their basic liberties actually do ignore them, with results that are salutary if anything. Hence, if Rawls is suggesting that it is morally wrong to take the law into one's own hands in cases of this sort, on the grounds that (i) society would be better off if we complied, or (ii) that society would be worse off if everyone acted as we do, he's just mistaken. I think the first assumption is obviously false and that the second too is, upon reflection, clearly mistaken. After all, most people do act as I do when similarly agitated by alleged injustice of the relevant sort.

But now consider cases of another sort: cases in which the law intrudes upon what a person believes to be a basic liberty, but where society (i.e., the majority of the voters or their representatives and/or the appropriate appellate courts) believes that the exercise of such a "right" would be detrimental to the rights or well-being of others. Here we would find laws against abortion, for example, laws against certain forms of racial discrimination, laws restricting sale and possession of certain firearms, and so forth.

The point of mentioning the second class of cases is that in many instances it's plausible to claim that no matter how sincere a person is in his belief that forced racial integration, for example, is unjust, society is justified in seeing to it that he is not allowed to act on his beliefs in ways that would be detrimental to the rights and well-being of his fellow citizens. Hence, we do not want to say that the sincere bigot is justified in breaking the laws against racial segregation, simply

because his bigotry is sincere and because he sincerely believes that it's unjust, for example, to bus his children from one school district to another. But on what principle can we allow "clear-headed" liberals to disregard the laws against smoking marijuana and against unnatural sexual congress, while depriving the muddle-headed bigot of the right to disregard laws that he feels just as strongly opposed to -- on grounds of their obvious injustice?

The suggestion here, of course, is that we face a problem of generalization: if we concede the moral propriety of private judgment in cases of the first sort, then we must make a similar concession in cases of the latter sort. And if we did this, the results would be disastrous. Or, more modestly, Rawls can simply claim that justice will be better served, in the long run, by disallowing private judgment in these cases than by allowing it.

My main complaint against Rawls here is that he makes no serious attempt to prove that this is so. I also suspect, what I cannot prove here, that Rawls is mistaken and that justice is in fact not better served by the kind of rigorous compliance with the principle of majority rule that he favors. It seems to me that we have to balance the beneficial effects of individual autonomy in cases of the first sort above against the possible ill-effects of private judgment in cases of the second sort. And of course, in order to do this, we would have to have some idea of the likelihood that people would use the principle of noncompliance which I favor just as often in cases of the second sort as in cases of the first sort. In other words, we would have to determine whether people would abuse or misuse my principle of private judgment as often as Rawls seems to think.

With respect to the question of weighing the pros and cons of extending the principle of private judgment in the fashion I favor, it is important to remember that Rawls himself believes that an injustice is tolerable (here: the injustice of preventing me from smoking marijuana or performing sodomy in the privacy of my own home, for example) only when it is necessary to prevent an even greater injustice. Hence Rawls

must believe that suppressing private judgment is to a large degree necessary in order to prevent some grave injustice in society. Since Rawls's criteria for social justice are the principles of justice, he must be claiming that a society which operated on some principle of "selective" compliance would tend to realize these principles less effectively than a society in which general compliance was the rule. My own view, on the other hand, is that the dangers of extending the principle of private judgment do not outweigh the dangers of majoritarian tyranny involved in Rawls's view and the subsequent loss of liberty that would accompany it. How does one settle a disagreement of this sort?

For one thing, we have allowed our discussion to become excessively abstract. Suppose we try to imagine what it would be like for one society to operate on a principle of "selective" compliance and for another society to operate on some more rigorous principle of general compliance. What's worrying Rawls here? At least two sorts of things, I think:

- (i) People would no longer feel morally bound to results of the principle of majority rule that they considered unjust.¹⁹
- (ii) People would be acting contrary to the principle of majority rule sufficiently often to make the society unstable -- and less effectively just than it might otherwise be.

I think that (i) is correct. Indeed, the point of my argument is that people are not -- and need not feel -- morally bound to results of the principle of majority rule that they consider unjust.²⁰ The second point assumes, I think, that people's conduct -- in complying or not complying with what they think are unjust laws -- is based on their moral commitment to the principle of majority rule. For remember: Rawls and I differ not on the question of whether unjust laws are laws or not, and hence have whatever force laws have in constraining people's behavior, but on the question of whether such laws have moral force as well as whatever other force laws have in a society like our own (e.g., the power simply to intimidate people). Hence I think that Rawls and I are arguing about an empirical question of the following sort: to what extent, under systems like our own, are people actually willing to comply with what

they believe are unjust laws simply because they are valid laws and because people generally recognize that they are morally bound to comply with valid laws, other things equal?

Personally, if I think a given law is unjust, then the only reason that moves me to obey it is the possibility of getting caught if I do not. But suppose Rawls is right. And suppose that people in a society like our own were suddenly to take their private judgments of the relative justice of duly enacted legislation more seriously than they do now and suppose that people subsequently began to disregard duly enacted legislation when they considered it unjust in the ways sketched above. What would be the likely consequences of this sort of situation?

There would, by hypothesis, be considerably more "conscientious" disobedience to the law than there is at present. People who were opposed to "busing" legislation, for example, would feel no moral compulsion to comply with it. But what's so troubling about this? Rawls's thought must be that in the long run this sort of conduct would lead to a situation of instability, as he sometimes puts it, or to a situation where people's basic political liberties and equality of opportunity were less effectively served than in a situation where they could count on each other to comply fairly rigorously with duly enacted legislation even when it was unjust.²¹

Unfortunately, not only does Rawls not present a shred of evidence in defense of these claims, but upon reflection they seem plainly false. As things are, it is clear that people disregard legislation of the sort in question when they can get away with it, and often even when they cannot. If people were to adopt my view, unless we changed our judicial system, they would still be accountable for disobeying the law, however pure their motives. The difference, however, is that while we would say that such people had acted illegally, we would not (with Rawls) say that they had acted wrongly or immorally.²²

Are we likely to face massive and debilitating disobedience on this hypothesis? Notice that we would be morally bound to Rawls's version of non-selective compliance only if we would face such problems.

But, then, to prove his moral claim Rawls needs to prove this empirical claim, something he hasn't done. What's more, it's hard to say how such a claim could be tested much less proven. In the next two chapters I hope to show that the most plausible defense of this empirical claim involves the additional claim that our courts could not possibly function if they were forced to hear pleas based on moral motivation of the sort we have been discussing (i.e., pleas based on the claim that a given law was unjust and that this justified noncompliance). I shall do my best to weaken the intuitive plausibility of this new claim and to outline a legal system that would accommodate conscientious disobedience to the law.

CHAPTER V

THE ARGUMENT BY ANALOGY

1. Rawls's argument that under certain circumstances we have a moral duty to comply with both just and unjust laws is both interesting and important in its own right. It defends explicitly and at some length a claim that is quite common in the recent literature on civil disobedience but that is seldom argued: the claim that there is a moral presumption in favor of obedience to the law, at least in reasonably just democratic societies, and that this presumption weighs heavily even when we are faced with an admittedly unjust law.¹

But this argument is of interest for other reasons as well. It is only in light of the contemporary strain of restricted or qualified "legalism" that this argument exemplifies, for example, that the intensive efforts of the past decade to show that civil disobedience is sometimes morally justified can be fully understood and appreciated.² What's more, two central contentions of Rawls's general theory of political obligation emerge in the analysis of this argument, contentions which seem to me to epitomize recent liberal thinking on this topic: (1) the claim that the obligation to obey the law is a moral obligation, one that extends as far as an obligation to obey unjust laws, under certain circumstances, and (2) the claim that at some point, even in a reasonably just democratic society, this obligation can break down and disobedience to the law can be morally justified.

As I have already indicated, Rawls's work seems to me to embody a further claim that has also been quite prominent in recent liberal politics and political theory: the claim that "paying the penalty" is an essential characteristic of justified civil disobedience. In this chapter I want to analyze the latter claim in relation to the first two. That is, I want to show first why it is tempting, given the tenets of liberal legalism, to make a willingness to suffer the consequences part

of the justification of civil disobedience in a constitutional democracy and, secondly, why the attempt to do so does not succeed.

In an important sense I shall merely be exploiting the argument of the preceding chapter and not considering it in its own right. For I shall assume that that argument can be made good. What I should like to investigate is the possibility that Rawls might be able to provide a defense of the willingness requirement that is analogous to (or in some sense related to) his defense of the claim that sometimes we are morally bound to obey unjust laws, at least in a reasonably just constitutional democracy.

Suppose, then, that Rawls's argument for restricted legalism with respect to legislative decisions can be made good; that is, suppose, among other things, that it is true that sometimes we have a moral duty to obey unjust laws. What follows -- either directly from this latter claim or by parity of reasoning -- with respect to Rawls's subsequent claim: that even when this duty to obey unjust laws is overridden, we should nonetheless willingly accept the consequences of disobedience in the kinds of cases we're concerned with?³ In other words, if we assume that a certain law is beyond the limits of "tolerable injustice", as Rawls puts it, and thus that civil disobedience is justified by Rawls's lights, and yet assume that the law itself and the society as a whole are not so unjust as to justify anything more than civil disobedience, why must we be willing to accept the consequences of disobeying the law or laws in question?⁴ In particular, how can the argument for restricted legalism with respect to legislative decisions be used to defend extreme legalism with respect to judicial decisions in a reasonably just constitutional democracy?

I must emphasize that the arguments for the willingness requirement that I am about to examine do not occur explicitly anywhere in Rawls's writings. It is not merely the principle of generosity, however, that induces me to introduce them on Rawls's behalf, although this is reason enough. As we saw in the preceding chapter, Rawls uses the same method for establishing the particular claims of his theories of political

obligation and civil disobedience that he uses for validating his general theory of justice (and his two "principles of justice" in particular). Hence it seems appropriate to ask whether this same method -- what can be called "the method of hypothetical agreement" -- can convince us that we should willingly accept the legal consequences of our illegal actions, even when the latter are admittedly justified.

2. Now it might be thought that Rawls does not actually have to argue for the willingness requirement independently of arguing for the claim that sometimes we have a moral duty to obey unjust laws: that he does not even need "the argument by analogy" that I shall outline below. For it might be contended that the willingness requirement is actually a logical consequence of the claim that in a reasonably just society we sometimes have a moral duty to obey unjust laws. This is what I shall call "the argument by implication", and I shall consider it briefly before getting into the argument by analogy, since if the former (stronger) argument is sound we don't need the latter (weaker) argument.⁵

The thrust of the argument by implication is roughly this. We have granted Rawls that we may indeed have a moral duty to obey unjust laws under certain circumstances. It would hardly make sense, however, at the stage of the hypothetical "constitutional convention", for example, to agree to obey certain unjust laws without also agreeing both to accept the consequences of disobedience when we don't obey such laws and to apply them when someone else doesn't obey. The basic idea here is a familiar one: given what we know about people, coercive laws without sanctions are about as good as no laws at all. So in recognizing that society will need laws, and in agreeing to obey them (even, up to a point, when they are unjust), we must also recognize the need to agree in advance to accept and apply the sanctions that make such laws effective (supposing, of course, that they do).

This is an important argument, even though it does not validate the willingness requirement. For one thing, it forces us to recognize the seriousness of just what we're getting into when we make the orig-

inal "agreement" (at the constitutional convention) to abide by at least some unjust laws: we must be willing not only to receive the whip's blows, as it were, should we break our agreement to obey such laws, but also to administer these blows when others fail to obey. But, by the nature of the case, this will be disobedience to laws which we admit are unjust but which are presumably not sufficiently unjust to warrant justified civil disobedience. To the conscientious man it might be even more repugnant to concede that society can be justified in punishing others for refusing to comply with unjust legislation than to agree that he himself will accept such punishment willingly.

Now it may be that Rawls can convince us that if we're going to agree that it's right to obey at least some unjust laws, then we must also grant that it's right to punish conscientious men for disobeying these laws -- as long as the laws in question do not exceed certain limits of injustice, of course. It's difficult to say whether this is simply an unfortunate consequence of the necessity of agreeing to obey some unjust laws, or whether it's a rather compelling reason for not making such an agreement in the first place (we are agreeing, after all, to punish the "innocent", or at least the "righteous", in certain circumstances). This is quite a serious dilemma, in my opinion, since it raises a difficulty which would take us back to the argument in Chapter IV and perhaps induce us to reconsider our concessive attitude towards that argument. We do not need to resolve this dilemma here, however, for the argument by implication does not validate the willingness requirement with respect to justified civil disobedience, which is our primary concern in the present chapter.

The most that this argument proves, even if it's basically sound, is that we have an obligation and/or a duty to accept (and apply) the consequences of disobeying even an unjust law in those cases where the injustice of the law was not sufficient to override our prima facie duty to obey such laws in the first place. But, of course, we're not concerned with cases of this sort here (although if they do imply that we ought to punish in cases of this sort, they raise serious problems for even these

cases, as I have just tried to show). We're concerned at present with those cases where we would all agree that the tolerable limits of injustice have been exceeded and hence civil disobedience is appropriate and morally justifiable. That is, we're concerned with those cases where our duty to act justly overrides our duty to obey the law. What's the argument for the claim that even in these cases we ought to accept the legal consequences of disobedience? It cannot be that we have agreed to obey such laws, and hence accept the consequences, for we haven't, even according to Rawls, agreed to obey them.

3. It is at this point that what I shall call "the argument by analogy" is relevant. Can we make a case for saying that while we agree to obey unjust laws, even in a reasonably just constitutional democracy, only up to a certain point, we agree to accept the legal consequences of disobeying unjust laws even in cases where this point has been passed and we no longer have a moral duty to obey them?⁶ If there is an argument for doing so, it will be a subtle one, even if it is analogous to the argument for agreeing to obey some unjust laws. For it will have to meet a compelling argument against punishing justified civil disobedience, what I shall call the argument that justification entails nonpunishment. Roughly, the latter argument is this:

- (1) When disobedience to a particular law (or laws) is justified, the prima facie duty to obey that law (those laws) is overridden.
- (2) When the prima facie duty to obey a particular law (or laws) is overridden, the disobedient should not be punished for violating the law (or laws) in question.
- (3) Therefore when disobedience to a particular law (or laws) is justified, the disobedient should not be punished for violating that law (those laws).

This argument is of special interest for two reasons. First, it is a general, theoretical argument against punishing justified civil disobedience. Thus far, we have only been considering arguments in favor of the willingness requirement and attempting to show that they are unsound.

But this procedure leaves open the possibility that there are still other arguments, not considered here, that will vindicate the willingness requirement. But if the argument that justification entails nonpunishment is sound, then we have a compelling reason for thinking that no argument that defends the willingness requirement, as a matter of moral and political principle, is sound. In other words, the argument above suggests that the willingness requirement is mistaken for general theoretical reasons and not simply because the particular arguments for it considered earlier are unconvincing.

But of course, we must ask just how convincing the argument that justification entails nonpunishment is. This brings us to the second point of interest regarding this argument. I think it can be shown that it derives whatever plausibility it has from Rawls's own theories of political obligation and civil disobedience. This becomes evident when we ask what reason there is for believing that premisses (1) and (2) above are true.

I can make my point here most simply by explicating (1) and (2) in a certain way, thus showing not only that they are quite plausible, upon reflection, but also that their plausibility is derived from some of Rawls's ideas in A Theory of Justice. Thus, suppose we ask what "justified" means in premiss (1). This premiss is certainly not obviously true on just any interpretation of "justified", for suppose that by "justified disobedience to the law" someone meant disobedience that is justified by one's own lights or justified on one's own private moral principles. Surely it would be at best controversial to claim that the prima facie duty to obey the law breaks down whenever it conflicts with one's private moral assessment of what one ought to do. However, if we explicate the notion of justified disobedience to the law as Rawls does in A Theory of Justice, and expand (1) accordingly, we find that this premiss is almost trivially true. Very roughly, the following is the sort of explication I have in mind:

- (1') When disobedience to a particular law (or laws) is justified by a clear and serious violation of the principles of justice, then

the prima facie duty to obey that law (those laws) is overridden because of the right to resist serious injustice (i.e., because of the intimate relationship, in Rawls's theory, between the principles of justice and the prima facie duty to obey the law).

Premiss (1') is really no more than a summary statement of Rawls's theory of the justification of civil disobedience. The idea, again very roughly, is this. Rawls develops both his theory of when we are morally bound to obey the law, as well as his theory of when we are morally justified in disobeying it, within the framework of his general theory of justice and his two principles of justice in particular. It is not surprising, then, that when we can say that a man is justified in disobeying the law, we can say also that his prima facie duty to obey the law has been overridden. Indeed, this is almost trivially true on Rawls's theory, since justified disobedience is defined in terms of a breakdown of the conditions that create a prima facie duty to obey. Hence, Rawls is not only stuck with (1'), it would seem to be axiomatic on his theory of political obligation. What about (2)?

This second premiss is curious. On the one hand, it seems to me to have a profound intuitive appeal, especially when we explicate it in light of (1'), our explication of the first premiss. This we can do as follows:

- (2') When the prima facie duty to obey a particular law (or laws) is overridden because of the right to resist serious injustice, the disobedient should not be punished for violating the law (or laws) in question because a man is liable to such punishment only when he actually had a duty to obey the law or laws in question.

If (2') can be made good, then we can infer (3'), which is simply a fuller version of (3), the original conclusion above:

- (3') When disobedience to a particular law (or laws) is justified by a clear and serious violation of the principles of justice, the disobedient should not be punished for violating the law (or laws) in question.

The last part of (2') is the crucial claim here; the key idea is

simply this: no duty, no punishment. That is, when the moral duty to obey the law breaks down, the moral right to punish disobedience (and the duty to accept it) breaks down as well. This seems quite plausible on its face, although there are a number of difficulties which could be raised about the intuitive idea behind this formulation. I shall not go into all these difficulties here, however, because Rawls is committed to denying (2') and because a very interesting argument can be constructed from his work for denying it; that is, for saying that justified disobedience to the law can be justifiably punished, even if we might not always choose to do so.

The argument I have in mind, which I shall eventually formalize and which I call the argument by analogy, runs roughly as follows: rationally self-interested men, at the stage of the hypothetical constitutional convention, would agree to accept the consequences of disobeying even those laws they are morally justified in disobeying (in light of the principles of justice) as long as the constitution and the society as a whole are still reasonably just; and such an agreement would be both reasonable and just. The last point is essential, for Rawls must show not only that such men would presumably make such an agreement but also that in doing so they would create moral ties that are binding at subsequent stages. Why would rationally self-interested men make such an agreement and why would they be acting rightly (as well as reasonably) in doing so?

To begin with, we must remember that when the question of whether or not to adopt some sort of willingness requirement arises, a number of "agreements" have already been made. The parties in question have already agreed, among other things, to comply with unjust legislation up to a certain point. They have also agreed, however, that at some point the obligation imposed by this agreement breaks down and civil disobedience can be justified, even though the constitution and the society as a whole are still reasonably just. In short, the parties have agreed to a general theory of political obligation and a particular theory of civil disobedience. But they have agreed, on a deeper level and at an earlier stage (in the "initial position"), to even more. They have agreed

that injustice is morally tolerable if (and only if) it is required to avoid even greater injustice (what I call "the principle of choice with respect to unjust alternatives"), and they have acknowledged a natural duty to support and foster reasonably just institutions ("the natural duty of justice", as Rawls puts it).

All of these previous "agreements", and the duties and obligations they impose, will be relevant to the question we now want to consider: would such men, and should they, morally speaking, make an absolute willingness to suffer the (legal) consequences part of their theory of justified civil disobedience?⁷ Starting from the point of view of men at the stage so far described, in other words, what can be said for and against agreeing to make the willingness requirement part of the theory of justified civil disobedience? Before constructing a formal analogue to the argument of the preceding chapter (see section 6 below), I should like to try to speak to these questions somewhat informally.

4. One obvious reason against making such an agreement, a reason embodied in the argument that justification entails nonpunishment, is that we are discussing cases where, ex hypothesi, the limits of tolerable justice have been exceeded (though not exceedingly so).⁸ In the absence of any other considerations, this seems like a good reason for not punishing (or willingly accepting punishment) in such cases.⁹ Why, despite this fact, would we agree to punish and be punished in such cases, and why would such an agreement be morally justified? An obvious answer suggests itself: ideally we wouldn't, and shouldn't (see the argument that justification entails nonpunishment), but we are unfortunately not dealing with an "ideal" situation. We are dealing with real men in a real society, and the problem of distinguishing cases of civil disobedience that are justified (in our theory) from those that are not is insurmountable (or so it might be argued). Let us call this the argument from "impracticability": it would simply be impracticable to try to identify (in a court of law, say) exactly those cases of civil disobedience that are justified, even on the universally accepted (Rawlsian) theory of civil disobedience, and to separate them from those that are

not. I shall speak to this difficulty shortly. But there is even more to the argument for the willingness requirement than this, and we should be clear about it from the start.

Let us grant for the moment that it would be very difficult, if not impracticable or even impossible, to distinguish (at the judicial level) cases of justified and unjustified civil disobedience. Nonetheless, this difficulty, if that's all it is, might be one that reasonable men would be willing to put up with for the sake of the greater justice of not punishing their fellowmen for disobeying intolerably unjust laws. But there is an obvious rejoinder that could be made to this point. For the claim behind the willingness requirement is not simply that we would encounter enormous (if not insurmountable) practical difficulties in trying to sort out, judicially, "justified" from "unjustified" acts of civil disobedience, but that if we allow ourselves (or our courts, rather) to undertake such a task, we will eventually, and perhaps very quickly, undermine the very stability of our otherwise "just" social system. And this of course is quite important, since we have already acknowledged that we have a natural duty to support and foster just institutions. What's more, we have agreed that since a certain amount of injustice is inevitable no matter what institutions we choose, we are justified in choosing (indeed, we are morally bound to choose) that arrangement which will produce the greatest balance of justice over injustice in the long run. That is, even an arrangement which involves us in a certain amount of injustice is morally tolerable, if it is required to prevent an even greater amount of injustice. I must say a bit more about these matters before going on.

It may appear to be obvious that it is better to have a viable and fairly stable society at the expense of systematically punishing justified civil disobedience than to have a "clean conscience" with respect to never punishing such disobedience at the expense of eventually undermining the social fabric of such a society. But aside from the difficulties involved in proving the claim that allowing justified civil disobedients to go free would involve us in an arrangement which would lead to substantial social

instability, difficulties which we shall consider in detail in the next chapter, Rawls is faced with an even greater difficulty here. For he has claimed repeatedly that one of the principal merits of his general theory of political obligation is that considerations of justice are absolute with respect to considerations of utility. Suppose we grant that a society without a willingness requirement would be less efficient and perhaps even less stable, in the long run, than a society with such a requirement. How can Rawls argue from this fact, if such it is, to the claim that we ought to choose the former sort of society over the latter without conceding that the efficiency and stability of the latter are sufficient reasons for legitimizing the prima facie injustice of punishing justified disobedience to the law? Doesn't Rawls have to qualify his claim that justice is "absolute" with respect to utility in order to get us to agree that we would, and should, punish each other for justified disobedience to the law?

I think there is no question that if Rawls's sole argument for the willingness requirement were that society would be less efficient without it, then his own convictions about the priority of justice over efficiency would force him to choose the more just but less efficient society over the less just but more efficient one. But surely it is incorrect to interpret Rawls as arguing for the willingness requirement simply on the basis of "efficiency" in some simple-minded utilitarian sense. His argument must certainly be that a society without a willingness requirement is less just than a society with one, and not simply less efficient. If he could prove this, then on his own principle that injustice is permissible when it is required to offset even greater injustice, we would be morally justified in choosing a society with a willingness requirement over one without such a requirement. Can Rawls prove this?

5. Everything depends here on how likely it is that the absence of a willingness requirement (or a similar constraint on illegal but morally justified actions) would lead, in an otherwise reasonably just society, to instability of the sort that would shift the balance of justice in an

undesirable direction. That is, the argument for the willingness requirement, as we shall see more perspicuously in a moment (section 6), depends on the claim that a greater balance of justice over injustice is achieved, in the long run, by a society that insists on a willingness to pay the penalty for otherwise justified civil disobedience than by a society that does not.¹⁰

If this claim is correct, then presumably our duty to support and foster just institutions would override our duty not to inflict suffering on someone for doing what was admittedly the right thing to do. The analogy to the situation in Chapter IV should be apparent: it would ordinarily be wrong to treat a certain group of people unfairly (as unjust laws sometimes require us to do), but the natural duty of justice can create an "actual" or "on balance" duty to do just this, as we saw in Chapter IV, if the stability of a just social scheme and the greater balance of justice require it. Similarly, although it is normally wrong to inflict suffering on someone (by way of judicial punishment and fines, as much as in any other way) for doing what we agree he ought to have done (even if this involves disobedience to the law), nonetheless, the natural duty of justice can justify such a practice if the stability of a just social scheme and the greater balance of justice require it.

In both cases we are faced with a straightforward but very complicated empirical question, which in the case at hand (whether or not to require a willingness to pay the penalty for justified acts of civil disobedience) takes the following form: what would be the empirical effects of the various competing arrangements? And just as in the argument for the duty to comply with unjust laws, as far as anything Rawls actually says is concerned, we are left to resolve this question on a very impressionistic, intuitive level. Rawls's intuitions seem to be that without a willingness requirement we would pretty quickly come to such acrimonious disagreements over particular cases (and perhaps over certain sorts of cases) that the social fabric in our generally just society would deteriorate and that the balance of justice would shift in an undesirable way.

This is a large claim, and it will be difficult to appraise it without a great deal of empirical information. Rawls acknowledges, with respect to a related problem, that "men of great honesty with full confidence in one another might make such a system work" -- that is, might make a system without a willingness requirement work; but he goes on to say that, nonetheless, "as things are, such a scheme would presumably be unstable even in a state of near justice."¹¹ But it is not enough to say (or even show) that such a scheme would be "unstable". The instability must, by Rawls's own lights, be of the kind that affects not only the efficiency of the social system but the long-run balance of justice in that society as well. Of course, these factors are likely to correlate highly in a given instance. We can summarize this by saying that, for Rawls, the optimal social arrangement is that which, among the possibilities available to us, is most effectively just. I shall sometimes use this last notion to gloss the intricate relationship between justice and efficiency in Rawls's conception of the just state. In short, Rawls's claim must be that a society with a willingness requirement is more effectively just, in the long run, than a society without one.

It goes without saying, I think, that the burden of proof is on Rawls to make this last claim good. In the previous chapter I conceded the analogous empirical claim that Rawls needed for his argument that sometimes we have a moral duty to comply with unjust laws. I wish to be more critical towards the analogue of that claim which Rawls needs for his argument that even when we are justified in disobeying an unjust law, we ought to pay the penalty for doing so. This new claim -- that an unwillingness to pay the penalty in such cases would be disruptive of the greater balance of justice -- is subject to criticism on at least two major counts.

To begin with, Rawls nowhere seriously considers the tenability of actually trying to implement a legal defense for acts of civil disobedience that the dissenter believes to be morally justified, so that there would be a public, nonviolent, institutionalized way of expressing (and defending) an unwillingness to pay the penalty for such conduct. No doubt such far-

reaching judicial reform would raise serious practical problems, problems we shall consider in the next chapter. And perhaps it is unfair to demand an account of these matters from Rawls, since it is not clear how strong his commitment to the willingness requirement is in the final analysis.¹² Nonetheless, to the extent that he is committed to it (see my discussion in Chapter II, section 2), the burden of proof is on Rawls to show that these practical problems are so great as to threaten social instability of just that kind which overthrows the optimal balance of justice in a democratic society. And in any case, arguments of this kind (arguments that justified civil disobedience can be justifiably punished) are common enough to merit our attention, even if Rawls himself would concede that they are not terribly persuasive.

There is a second preliminary point that Rawls's critic might make with respect to his apparently strong stand on paying the penalty for justified civil disobedience. This point arises out of a curious feature of Rawls's position on the issues of obeying unjust laws up to a point, on the one hand, and accepting the consequences of disobedience even when the injustice of the law is beyond that point, on the other. Rawls's position on the duty to obey unjust laws is not absolute or unqualified; that is, he adopts (and defends) what I call "restricted" rather than "extreme" legalism with respect to legislative decisions. And yet his position with respect to judicial decisions -- like Socrates' position in the Crito -- is more or less absolute or unqualified.¹³ That is, he adopts a position of "extreme" or "unrestricted" legalism rather than "restricted" legalism with respect to judicial decisions, at least in reasonably just constitutional democracies which are at a point where no more than civil disobedience is justified on my "circumstances continuum" of Chapter II.

This is a curious anomaly, especially in a thinker as careful as Rawls. Either Rawls thinks that compliance with judicial decisions needs even more stringent safeguards (in a reasonably just democratic society) than compliance with legislative decisions, or he has not followed the logic of his own restricted legalism to its ultimate conclusion. I should

like to show now that the reasons for safeguarding judicial procedures in a more or less just democratic society are indeed compelling ones; I should also like to show, however, that we can safely protect such procedures without resorting to legal extremism of the sort embodied in the willingness requirement. To do this I shall have to construct a formal argument for the willingness requirement that is exactly analogous to the argument for the duty to comply with unjust laws. Not even this final argument will do the job for Rawls, as I shall show. But it is worth constructing, and not simply because only by doing so can I make good on my claims against the willingness requirement. It turns out that in considering the formal analogue to the argument sketched in the last chapter, some extremely important points with respect to my own conception of the place of punishment in the theory of civil disobedience come out.

6. At the end of Chapter II I considered, very briefly and sketchily, an argument of Professor Charles Fried's for something very much like the willingness requirement. This argument was based on what Fried calls "the principle of institutional settlement". It turns out that this principle, or some variant of it, is crucial to the formal version of the argument by analogy that can be developed from some of Rawls's work.

Fried's principle of institutional settlement can be interpreted as the judicial analogue of the principle of majority rule.¹⁴ Thus, on this reading, in agreeing to the principle of institutional settlement, we are agreeing to abide by the forms and consequences of trial by jury, for example, just as in agreeing to the principle of majority rule we are agreeing to abide by the majority's decisions (only up to a point, significantly) on legislative matters.

I believe that interpreting the argument for the willingness requirement in terms of this analogy brings out what is at the heart of the commitment to various versions of the willingness requirement in the recent literature on the justification of civil disobedience. But this will be obvious only in the sequel. The principle of institutional settlement,

qua analogue of the principle of majority rule, must first be worked into a formal argument for the willingness requirement. This can be done in the following way:

- (1) A society is reasonably just if its basic structure more or less satisfies the principles of justice.
- (2) If a society is reasonably just, then we have a natural duty to support and comply with its institutions if it exists and applies to us.
- (3) Therefore, we have a natural duty to support and comply with the institutions of a reasonably just society when we find ourselves in such a society.
- (4) Supporting the institutions of a reasonably just society involves supporting and complying with one of its most fundamental and essential institutions -- a just constitution.
- (5) Therefore, when we find ourselves in a reasonably just society, we have a duty to support and comply with its constitution (as long as both the society and the constitution are just -- or as just as can reasonably be expected).
- (6) Supporting and complying with a just constitution in a reasonably just society requires supporting and complying with the principle of institutional settlement with respect to judicial matters -- i.e., the judicial system as we know it (more or less).
- (7) Therefore, we have a duty to support and comply with the principle of institutional settlement with respect to judicial matters in a reasonably just society.
- (8) Supporting and complying with the principle of institutional settlement with respect to judicial matters requires accepting (and applying) the standard legal penalties for disobedience to the law even when that disobedience was morally justified, justified either because the law itself was beyond the limits of tolerable injustice or for some related reason articulated by the theory of justified civil disobedience.
- (9) Therefore, in a reasonably just society with

a just constitution we have a duty to accept (and apply) the standard legal penalties for disobedience to the law, even when that disobedience was morally justified.

The crucial premisses here are (6) and (8), of course, since the other premisses are more or less identical with premisses of the argument in the preceding chapter, premisses which we have agreed to grant for the sake of argument. Premiss (6) sets us up for, and premiss (8) ensures, the willingness requirement as we have been conceiving it. What's the argument for (6)? Roughly, and by analogy to the argument of the preceding chapter, we may suppose that just as we would agree to the principle of majority rule as a procedural device for resolving conflicts about legislative matters, so we would agree to the principle of institutional settlement with respect to judicial matters for resolving similar disputes on the judicial level. That is, the principle of institutional settlement is to be thought of as the most efficient way of handling disputes about how a given law is to be interpreted and applied, once it has been passed, just as the principle of majority rule is apparently the most efficient way of handling disputes about what the law is going to be in the first place.

Of course, there's much more to be said for these principles -- the principle of majority rule and the principle of institutional settlement with respect to judicial matters -- than that they are the most efficient ways of handling these difficulties. Indeed, more important, for thinkers like Rawls, than their bare efficiency is the fact that they are presumably the most effectively just ways of handling the difficulties in question. It is significant that from the point of view of justice it seems that, if anything, more could be said in favor of the principle of institutional settlement with respect to judicial matters than can be said for the principle of majority rule with respect to legislative matters. And in any case, it seems at least as reasonable to agree in advance (at the stage of the hypothetical "constitutional convention", say) to abide by the principle of institutional settlement when

judicial disputes arise as it is to agree in advance to abide by the principle of majority rule when legislative disputes arise.

Now premiss (6) can either be read as making a fairly noncontroversial claim about the reasonableness of adopting a certain kind of institutionalized procedure, in which case the burden of subsequent argument is thrown on (8), or it can be read as making a fairly strong claim, in which case we shall want to ask some serious questions about (6) as well as (8). I wish to avoid the latter difficulty by reading (6) in such a way that it is relatively noncontroversial. In other words, I want to see how much weight premiss (8) can bear. This should not be taken as implying that (6) is altogether empty of substantive content. On the contrary. In committing ourselves to the principle of institutional settlement with respect to judicial matters, we are presumably committing ourselves to a great deal. Let us suppose that in a very general way this amounts to a commitment to something like our own American judicial system. I say "in a very general way" because we want to leave open the possibility that we have not committed ourselves to exactly these procedures. This is the point of saying that (6) is not intended to carry much of the burden of argument. It merely stipulates that we have agreed to trial by jury, for example, and some of the other general features of our own judicial system. What we should like to know is how such a commitment bears on the theory of justified civil disobedience. In particular, we should like to know if a commitment to the principle of institutional settlement necessarily commits us to willingly paying the penalty for disobedience that is admittedly justified from the moral point of view, as (8) contends. What is the argument for (8)?

We must not expect to find any explicit arguments for (8) in Rawls's work, or anyone else's for that matter, since (8) is an artifice of my own. However, as I shall show, when we try to imagine how the defense of (8) would proceed, we quickly see what the ultimate rationale behind the willingness requirement really is. Very roughly, (8), and the willingness requirement it entails, rests on the claim that we simply cannot settle moral and political questions in courts of law.¹⁵ Thus, the argu-

ment for (8) rests on both normative and empirical claims. The relevant normative claims are embodied in the so-called "natural duty of justice" (the duty to foster and support just institutions) and the "principle of choice", as I call it, with respect to unjust alternatives: when faced with a necessary choice among alternatives each of which involves a certain amount of injustice, choose that which will bring about the least injustice. The crucial empirical claim, of course, is that a system where civil disobedience is regulated by a willingness requirement of some sort is more likely to produce a greater balance of justice over injustice in the long run than a system where this is not the case.

Now we have encountered these normative claims before (see Chapter IV) and once again I wish to concede them for the sake of argument. My principal question is about the empirical claim that (8) needs to make in order for these normative premisses to yield the appropriate conclusion (that even when justified in his civil disobedience, a dissenter should pay the legal penalty willingly). But before asking how plausible the empirical claim is, we ought to get clear on just what it amounts to.

Premiss (8) above can be made good, as we have seen, only if it is true that a social system that requires a willingness to pay the penalty for justified civil disobedience produces a greater balance of justice, in the long run, than a system that does not. This in turn depends upon the claim that, among the alternative institutions available for consideration at the hypothetical "constitutional convention", the system with the willingness requirement is morally preferable to one without it because of the greater balance of justice adhering to the former alternative.

Now we have a rough idea of what the first alternative is: it is a system much like that defended and practised in the past ten years by most theorists and practitioners of civil disobedience, a system where one breaks the law publicly and nonviolently and then willingly pays the penalty. What is the second alternative, however, over which the former is supposed to win in the "balance-of-justice" calculation? Remember

that we have already conceded that in the sorts of situations we are concerned with, an unwillingness to pay the penalty cannot be exhibited by violent or secretive means. That is, we have agreed to abide by the requirements of publicity and nonviolence in the definition/justification of civil disobedience. So the alternative to the system with a willingness requirement, that is, a system where an individual can be justified in his disobedience without exhibiting a willingness to pay the penalty, must be a system where this unwillingness is expressed through neither violent nor clandestine acts.

I submit that the alternative we're describing is a society whose judicial system includes explicit legal defenses for civil disobedience based on the claim that the disobedience was morally right or even morally necessary. For reasons I shall explain in Chapter VI, I shall refer to such defenses summarily (and perhaps somewhat inaccurately) as defenses of "moral necessity". The idea, very roughly, is that such a system would allow the accused in a court of law to concede that he broke the law in question but to go on to defend himself for doing so by explaining why he feels he was morally justified in doing so -- for example, because his disobedience would be justified on Rawls's theory of civil disobedience.

It is this sort of judicial system, I think, where pleas of "moral necessity" are allowed to be entered in defense of a charge of criminal activity like civil disobedience, that we must take to be the principal alternative arrangement to the system Rawls chooses -- the system embodying the willingness requirement and prohibiting defenses of this sort in the courtroom. Hence, the crucial empirical claim on which premiss (8) above rests comes down to this: a system that prohibits (for the most part, if not entirely) legal defenses aimed at showing that the disobedient was morally justified in an admittedly illegal action tends to produce a greater balance of justice over injustice (in the long run) than the alternative system in which such defenses, of justification due to "moral necessity", say, are allowed. And it is this claim whose plausibility we must now examine. In order to do so, we shall have to get clear

on what the alternative arrangement I favor would involve, what can be said for and against such an arrangement, and whether the considerations against it outweigh those in its favor. Our ultimate criterion in resolving the last question of course is this: which arrangement is more effectively just or produces the greater balance of justice over injustice in the long run?

CHAPTER VI

JUDICIAL IMPLEMENTATION: SOME TENTATIVE SUGGESTIONS

1. Nearly a decade ago Professor Richard Wasserstrom wrote that it is "more than a little surprising...that the task of examining the possible procedures by which a legal system either could or should decide cases has on the whole been neglected by legal philosophers."¹ In this chapter I should like to suggest and assess one substantive procedural change that might be made in our own legal system, a change which would allow a legal defense for conscientious disobedience to the law. This defense, which I shall describe in detail in section 4 below, is to be thought of as specifying an alternative arrangement to the one Rawls seems to favor and which is in fact the arrangement favored by most recent theorists and practitioners of civil disobedience.

Rawls and other proponents of the willingness requirement need to defend some version of premiss (8) in the argument of the preceding chapter. I intend to show that (8) rests on a dubious empirical pre-supposition -- on the assumption that a legal system with a willingness requirement tends to produce a greater balance of justice over injustice than a system without such a requirement -- and I shall do this by showing that (8) rests on the mistaken idea that it is not feasible to allow moral considerations to be entered as part of a legal defense for disobedience to the law.

The basic thrust behind my own proposal is this. If an individual feels that by complying with a certain law he will be either the agent or the victim of serious injustice,² then (i) it is possible that he would be morally justified in refusing to comply with that law, and (ii) there is no good reason for refusing to give him a chance to defend his noncompliance in a court of law and to be released from the ordinary sanctions for noncompliance if he can convince a jury that his moral

beliefs are sincere and reasonable ones.³

It is obvious that no one who believes that disobedience to the law can never be justified, at least in a reasonably just constitutional democracy, will be receptive to my view. But the thinkers we have been concerned with in the preceding chapters do believe that disobedience to the law can be morally justified, even in more or less just democratic societies. And yet these same thinkers believe that even when disobedience to the law is justified, the disobedient must be ready to accept the legal consequences of his illegal dissent. My guess is that, ultimately, they hold the latter view because they hold some version of premiss (8) in the argument by analogy. This means that they oppose my view, if they do, because they disagree with the second part of it, (ii) above.

What sorts of reasons are there for denying a conscientious disobedient an opportunity to state his case in court? That is, what can be said against the notion of allowing someone to claim in court not that he did not disobey the law, but that he did and was morally justified in doing so? In trying to come to grips with the issues that surround this question, I shall limit my discussion to a rather narrow range of cases of conscientious disobedience. This limitation will obviously set limits on what I can expect to prove, but this is inevitable in a study of this sort.

In order to illustrate the kinds of cases I shall be concerned with, I want to return to Rawls's work on justified disobedience to the law. Rawls describes two forms of disobedience in his general theory of non-compliance: civil disobedience and conscientious refusal. Although he is willing to acknowledge that disobedience of either sort can be justified -- indeed, he tries to outline the conditions under which this would be so -- he also believes that the civil disobedient ought to accept the legal consequences of his illegal behavior. Rawls does not take a stand with respect to paying the penalty in cases of conscientious refusal. But if I am correct in speculating that he accepts some form of premiss (8) in the argument by analogy, then he would be implicitly committed to

saying that the dissenter ought to be willing to pay the penalty in both sorts of cases, unless express legal provision has been made for noncompliance, as in the case of conscientious objectors under the Selective Service System.

In any case, there are plenty of people who do hold that even when conscientious refusal is morally justified, this feature of the act is irrelevant from the law's point of view.⁴ This general view is in fact enshrined in our own legal system. If a young man opposed the war in Vietnam on the grounds that it was an unjust war or a war that was blatantly violating the principles of just warfare, and if on these grounds he refused induction into the armed forces, he was not allowed to defend this conduct in terms of his moral convictions about the war.⁵ Similarly, if women, or men over thirty, who could not oppose the war by refusing induction, expressed their opposition indirectly by violating some other law, their moral reasons for doing so were not the sort of considerations that would be permitted as evidence on their behalf in a court of law.⁶ These, and cases like them, are the kinds of cases I shall be concerned with below.

If we believe that at times civil disobedience and conscientious refusal can be morally justified in cases like these, cases of allegedly intolerable injustice, then why do we refuse to allow the considerations that motivate such noncompliance to be used in defense of it in our law courts? Suppose that we were to allow some sort of legal defense for conscientious disobedience to the law. In this case it would no longer be axiomatic that when one disobeyed the law on conscientious grounds one would have to be ready to pay the penalty without reservation. Of course, neither would it be axiomatic that one would never have to pay the penalty for conscientious noncompliance. Rather, whether or not one finally paid the legal penalty would depend on whether one could convince a jury of his peers that the grounds of his noncompliance were good grounds. But what would count as "good grounds" in cases of this sort?

In answering this question I plan to make things easy for myself, at least in one respect. Rawls has given us a fairly comprehensive account

of what constitutes good grounds for civil disobedience and conscientious refusal, and I propose to accept his account. That is, I shall say that civil disobedience is morally justified when the dissenter faced a serious violation of the first principle of justice or a blatant violation of the second part of the second, when he has tried the normal channels of legal redress without success, and when there is no reason to believe that others, similarly situated, will express their equal right to dissent, with harmful consequences for society as a whole. I shall say that conscientious refusal is justified, following Rawls, when a nation violates its jus ad bellum (wages an unjust war) or its jus in bello (conducts unjust warfare in an arguably just war). The question we are then faced with is simply this: if people can be morally justified in disobeying the law on grounds like these, why cannot the same grounds be used in a law court as part of a standard legal defense for conscientious noncompliance?⁷

2. The recent literature has not been entirely without advocates for some form of accommodation for conscientious disobedience to the law. I can set my own proposal in a clearer perspective by discussing a few of these proposals first. For it seems to me that one of them is clearly too weak to meet our needs, while the other is arguably too strong (or at least too sweeping).

In an article entitled "On Not Prosecuting Civil Disobedience",⁸ Ronald Dworkin suggests an interesting way of accommodating some forms of conscientious disobedience to the law. He argues that federal prosecutors should use the discretion they have for not prosecuting in cases of the sort mentioned above -- cases of conscientious resistance to the draft laws, for example, because of the war in Vietnam. Dworkin points out that although society "cannot endure" if it tolerates all disobedience to the law, it obviously does not follow "that it will collapse if it tolerates some."⁹ He then catalogues some of the many cases in which prosecutors can (and often do) decide not to press charges in criminal cases, and argues that there are "some good reasons for not prosecuting those who disobey the draft laws out of conscience:

One is the obvious reason that they act out of better motives than those who break the law out of greed or a desire to subvert the government. Another is the practical reason that our society suffers a loss if it punishes a group that includes...some of its most thoughtful and loyal citizens. Jailing such men solidifies their alienation from society, and alienates many like them who are deterred by the threat.¹⁰

The last few sentences suggest that Dworkin would appeal to both utility and to some principle of "fittingness" or equity as reasons for not prosecuting such disobedience.¹¹ These are reasons to which I shall also appeal in support of my own proposal for accommodating conscientious disobedience to the law. But Dworkin's suggestion, and others like it, while in the spirit of the suggestions I shall make below, are nonetheless far more limited proposals than my own. To begin with, Dworkin is speaking directly only to conscientious disobedience inspired by the war in Vietnam; he is speaking only indirectly, if at all, to the issue of how we might handle conscientious disobedience generally.

More important, however, is a limitation arising from the fact that Dworkin is suggesting discretion at the prosecutorial level. This raises two quite different problems, from the point of view of the present study.¹² On the one hand, looking at the problem from a purely practical point of view, it seems to me that it is unreasonable to expect federal prosecutors to take Dworkin's admonitions to heart and to refrain from prosecuting in the relevant cases. These are men whose careers depend, in part, on successfully prosecuting a certain number of cases, as well as certain kinds of cases. What's more, these men are representing in the courts the very governmental forces the resisters oppose.

Perhaps Dworkin has reasons for being optimistic in this regard, however. For example, perhaps when he made his proposal he believed that the prosecutors themselves recognized that more was to be gained by avoiding courtroom confrontations on the moral issues surrounding the war than by promoting them.¹³ But there is another more serious

difficulty connected with proposals like Dworkin's. For even if it is not unreasonable to expect federal prosecutors to act on his advice, an assumption that is controversial at best, it seems both dangerous and unnecessary to leave the discretion to decide on the fate of such dissenters solely in the hands of a class of men less likely than most to be sympathetic to the dissenter's claims.¹⁴ To be sure, it might often be best, for all parties concerned, to avoid a direct confrontation with respect to the moral issues at stake simply by not making an issue of them. But to leave the matter there overlooks the possibility that some dissenters may actually want to bring their cases to court as well as the probability that others will be forced to do so, whether they want to or not, by prosecutors who do not decide against prosecution. Both groups, I shall argue, ought to be able to avail themselves of a legal defense that allows them to defend their illegal conduct on moral grounds.¹⁵

Thus far I have been raising more or less practical problems that Dworkin's proposal must face, especially the problem of just how likely it is that federal prosecutors will in fact exercise the discretion that Dworkin is exhorting them to use. Before showing that suggestions like Dworkin's face even more serious theoretical problems -- problems of moral and political principle -- it might be worthwhile to connect Dworkin's remarks with a more general trend in legal thought to which they are related. This is the view that while moral issues are irrelevant to the court in determining the legal guilt or innocence of an accused draft resister, say, such issues are certainly not irrelevant at the level of sentencing.¹⁶ A conscientious judge, in other words, will take the moral integrity of the defendant into account when sentencing him, even though he cannot (and should not) do so when trying him for the crime. Unfortunately, this move is open to the same objections as Dworkin's view.

There is, to begin with, the problem of determining just how likely it is that a federal judge will be moved to leniency in the sorts of cases we are discussing. What's more, it should be apparent that whether pros-

ecutor or judge is thought of as exercising the discretion we are discussing, it seems very odd to leave decisions of this magnitude to individuals who are not especially likely to represent the moral consciousness of the community at large and who are in fact more likely than not to represent a segment of the community with very special (and very conservative) views about the propriety of any disobedience to the law, whether conscientious or not.¹⁷ As we shall see below, there is good reason to believe that one of the most important characteristics of the notion of a trial by a jury of one's peers is the fact that a carefully selected sample of a man's peers is more likely to be in touch with and representative of contemporary moral thinking than law enforcement agents like judges and prosecutors. If this is true, it would be salutary to allow a jury to hear and deliberate upon the relevance of a man's moral beliefs to the question of his liability to conviction and punishment for an otherwise illegal act.

We are now dealing with a more general view than Dworkin's view that the moral principles of war resisters, say, might appropriately be taken into account at the prosecutorial level. We are considering the view that such considerations can be taken into account both at this level and at subsequent levels when fines and punishment are being considered, supposing that prosecution has been carried through and carried through successfully. In other words, we are considering the not uncommon view that Dworkin's proposal can be supplemented by the proposal that we can (and should) respect the conscientiousness of principled disobedience to the law by urging the judge, for example, to mitigate the sentence for conduct of this sort, or even by urging the prosecution to ask for a mitigated sentence once a conviction has been achieved. I shall not belabor the practical difficulties that viewing the matter in this way seems to me to face, although I think they are formidable (see my remarks above).

But suppose I am mistaken or suppose I have exaggerated the extent to which difficulties like those suggested above are exigent. Nonetheless, it seems to me that a proposal such as Dworkin's faces serious

theoretical problems as well -- problems of moral and political principle, that is. And, more importantly, it seems that these same difficulties are attached to the more general view just sketched, the view wherein we would respect the conscientiousness of principled disobedience to the law by mitigating or even suspending sentence after conviction on standard legal grounds.

The claim of these dissenters is not simply that they are conscientious in their resistance, but that they are conscientious and correct. That is, at least in the cases we're concerned with (where the disobedience is justified ex hypothesi), the dissenters correctly believe that they have acted on moral principles that override their prima facie duty to obey the law and hence that the government is acting in ways that forfeit, at least to a certain extent and with respect to certain laws, its right to expect compliance.

Now when the court mitigates a sentence for criminal conduct, and even when the prosecution decides not to press charges in the first place, or to ask for a mitigated sentence once conviction has been achieved, it seems to me that the dissenter's moral stance is not receiving the sort of response it deserves (and implicitly demands) or, at best, is receiving such a response in an inappropriately oblique manner. When we mitigate a sentence, for example, we are in effect saying that while there has been an unjustified criminal act, we are willing to refrain from punishment, or to mitigate it to a certain extent, for various reasons. But is this what we want (and ought) to say to war resisters (and civil disobedients generally) whom we believe are justified in their disobedience to the law? I should think not. We ought to congratulate such dissenters for their courageous actions; and if we believe they really were justified in what they did, I should think we would want to say they are not deserving of punishment at all.

3. At this point we are approaching the basic question of why the moral considerations that motivate a conscientious disobedient's conduct should be deemed appropriate at certain points in the judicial process but not at

others. In particular, why should such considerations be irrelevant to a determination of guilt or innocence in the trial itself? Professor Jon Van Dyke has addressed this question directly in a recent article on "The Jury as a Political Institution".¹⁸

Van Dyke opens his article by pointing out that at the start of the trial of the "Chicago Eight" for allegedly conspiring to produce violence at the 1968 Democratic Convention, Federal District Judge Julius J. Hoffman "told the jurors that they must always follow his instructions on matters of law".¹⁹ According to Van Dyke, defense lawyer Leonard Wein-
glas objected, saying:

The defense will contend that the jury is a representative of the moral conscience of the community. If there is a conflict between the judge's instructions and that conscience, it should obey the latter.²⁰

Van Dyke comments dryly: "Judge Hoffman overruled the objection and the trial proceeded." The rest of Van Dyke's paper can be seen, I think, as an attempt to show that Hoffman's ruling was incorrect. He writes:

Although the American jury is still praised as a bastion of democracy, standing between oppressive governments and the people, most of today's American judges in fact do everything they can to emasculate the jury until the only role left for them is to review the facts and then rubber-stamp the application of the law for the government.²¹

Van Dyke observes that, according to a study made by Professors Harry Kalven and Hans Zeisel (THE AMERICAN JURY), "the most important reason jurors rarely use their theoretical power [to acquit a guilty defendant, for example, because he acted conscientiously] is that they are constantly told that they have none."²² A typical example, according to Van Dyke, occurred in the case of Dr. Benjamin Spock, when the jury returned a verdict of guilty for conspiracy to violate the draft laws. Van Dyke points out, as is well-known from subsequent press stories, that one juror said after the trial that he was "uneasy" about convicting the defendants, and yet, according to Van Dyke, felt that "he had no choice but to follow the judge's instruction." "The paradox," this juror went on,

"was that I agreed wholeheartedly with these defendants but...I felt that technically they did break the law."²³ It is apropos this remark that Van Dyke states his central thesis:

I contend that justice would be better served if jurors were told they have the power to act mercifully toward the defendant should they decide that applying the law to this act would lead to an unjust result.²⁴

He traces the history of judicial opinion against this stand from the early nineteenth century to 1895, when the Supreme Court of the United States considered the question in the case of *Sparf & Hansen v. The United States*:

The majority opinion of the Supreme Court spends forty-two pages reviewing the earlier decisions and comes to the conclusion that because we cannot tolerate allowing jurors to increase the penalties or make law on their own, we cannot tolerate allowing them to reduce the penalties or nullify laws.... Even though no one in the case before the Court argued that the jury should be allowed to create its own crimes or to render stiffer punishment than the law allows, the Court was haunted by that specter. Because the Supreme Court could not distinguish between lowering and raising the punishment it deprived the jurors of the power to do either.²⁵

Now Van Dyke makes so much of this distinction between lowering the punishment (or acting mercifully) and raising the punishment (or acting harshly) because he believes that in light of it the Supreme Court's position will be obviously untenable. He returns to the trial of the "Chicago Eight" and writes:

Leonard Weinglas was arguing to Judge Hoffman only that the jurors have the power to temper the law with mercy. He was not saying that the jurors had the power to interpret the law for any purpose beyond the confines of the courtroom. He was certainly not arguing that the jury should be empowered to hand down a harsher punishment than permitted under current law.²⁶

Similarly, he distinguishes between jurors making law and nullifying it,

because he favors the latter and not the former, and because he believes that the position against so-called "jury nullification" is based on a fear of the former and simply fails to make an obvious distinction. Thus, he writes:

Proponents of jury nullification are not arguing that the jury be given free rein to create new laws or define 'people's crimes' in a fashion reminiscent of Nazi Germany. The argument today is only that the jury has the right to mitigate existing laws and that this is a basic safeguard against an oppressive or even merely overly aggressive government.

Our judicial system already recognizes to some extent a jury acting vengefully and one acting mercifully. Appellate judges can reduce a sentence or refuse to uphold a conviction if they feel a jury has imposed too harsh a judgment on the accused, but they are prevented from reviewing a judgment of acquittal no matter how irrational it seems.²⁷

The position that Van Dyke himself supports, however, is both too vague and too wide-ranging. He argues that jurors should not be instructed that it is their duty to follow the letter of the law as the judge interprets it, but that they should be encouraged "to give respectful attention to the laws", on the one hand, and then instructed that nonetheless "they have the final authority to decide whether or not to apply a given law to the acts of the defendant on trial before them."²⁸

What we need to know, however, is on exactly what kinds of grounds a jury ought to base this sort of decision. Van Dyke believes that "jurors should be told that they represent their communities" and told as well that it is appropriate for them "to bring into their deliberations the feelings of the community and their own feelings as well."²⁹ It seems to me that talk of "representing one's community" and bringing one's own (as well as one's community's) "feelings" into one's deliberations in a criminal trial is simply too vague and too broad a charge to give a jury. It is too vague because one would like the jury to be fairly clear about what counts as a morally defensible (if not altogether correct) position as opposed to an indefensible (or downright incorrect)

moral position. Van Dyke's suggestion is too broad, on the other hand, insofar as it seems that moral considerations are at least prima facie relevant in any sort of criminal case on his view. I shall argue, to the contrary, that only very specific moral beliefs should be arguable in court and only in a carefully circumscribed range of cases (see section 4 below).

It is only fair to Van Dyke to note that at one point he comes somewhat closer to articulating the kind of moral standards he has in mind (though he still leaves open the range of cases in which they might be applied). He argues that the jury should be told that, despite their respect for the law, nothing prohibits them from acquitting the defendant "if they feel that the law as applied to the factual situation before them would produce an inequitable or unjust result."³⁰ This is still rather vague talk, but it raises the issue of justice and injustice that lies at the heart of recent concern about how the law can go wrong and how illegal dissent can be morally justified (and sometimes even morally obligatory). In the next section I shall try to delineate with considerably more precision what the relevant considerations of justice and equity are, the scope of their relevance in an ideal judicial system, and precisely how the arrangement I favor might be worked into a legal system like our own.

4. My own proposal differs substantially from Dworkin's, as well as the more general line on mitigation-after-conviction. I wish to argue that certain moral considerations should be allowed as part of an explicit legal defense that would enable a jury to acquit a defendant on the grounds that he was morally justified in what he did (or that he was in fact morally bound to act as he did) and hence that he cannot be held legally culpable for not refraining from an otherwise illegal act. This proposal also differs from Van Dyke's in a number of important ways. To begin with, I shall attempt to make the mechanics of the defense fairly specific. What's more, my suggestion is not that juries be given a general power to nullify law when they see fit, but that the defendant be

given the specific right to enter a plea of "moral justification". This plea will be based on a very explicit conception of what it is for disobedience to the law to be morally justified: roughly, the conception advanced by Rawls in A Theory of Justice, minus the willingness requirement. On this view the jury would be empowered to acquit if the defendant makes good on his claim that he was morally justified in disobeying the law. Obviously, any reasonable arrangement for allowing pleas of moral justification will have to be very specific not only with respect to what will count as a convincing moral defense but also with respect to how such a defense would actually be implemented in the criminal and civil law. I turn to these matters now.

In section 3 of the preceding chapter I tried to show that the argument that justification entails punishment gains a great deal of its plausibility from Rawls's arguments in A Theory of Justice. Premiss (1) in that argument, at any rate, according to which disobedience to the law is justified only when the prima facie duty to obey the law is overridden, is simply a gloss on Rawls's theory of the justification of civil disobedience, as well as his theory of justified conscientious refusal. According to both theories, there is a point when the limits of tolerable injustice have been exceeded and when the right to resist injustice overrides the duty to obey the law. Rawls delineates the conditions that indicate when this point has been reached rather precisely. To be sure, Rawls does not give us a simple and easily applicable set of necessary and sufficient conditions for the justification of disobedience to the law. But his theories do indicate rather nicely the sorts of cases that would qualify as more or less clear cases of justified illegal dissent and the sorts of cases that would certainly not qualify. Henceforth I shall assume for the sake of argument that the standards for when illegal dissent is morally justified are roughly the standards outlined by Rawls in his book.

Rawls's theory leaves a great deal of room for controversy, but from our present perspective it is important to note that it also provides

grounds for a considerable amount of substantive agreement on an impressive range of cases. This is because, like his theory of justice generally, Rawls's theory of noncompliance is designed to suggest the sorts of standards which rationally self-interested parties would be likely to agree to employ for resolving competing claims -- in this case, claims about whether a given law or policy had exceeded the limits of tolerable injustice and, hence, whether disobedience would be justified, all things considered.

In light of all this, it is possible to sketch roughly what a defense like the one I advocate would involve. Essentially, in allowing a legal defense of "moral justification" we would be allowing an appeal to Rawls's principles of justice or to what Marshall Cohen has called "the principles of political morality".³¹ More specifically, we would be establishing an arrangement where the question of whether or not an agent was justified in his disobedience to the law is to be settled by asking whether that disobedience would be justified on Rawls's theory of civil disobedience or on his theory of conscientious refusal. Thus, a jury would be faced with the problem of determining whether a disobedient's act was justified by a serious violation of Rawls's first principle of justice (the principle of maximal equal liberty) or the second part of the second principle (the principle of open offices).³² Alternatively, in certain cases the jury would be faced with deciding whether illegal dissent was justified on the grounds that a nation was violating what Rawls calls its jus ad bellum (waging an unjust war) or its jus in bello (by engaging in unjust warfare). A great deal of illegal dissent in the civil-rights movement in the 1960s was based on principles of the first sort. And most of the illegal dissent in the recent anti-war movement was based on principles of the second sort.

Now there are at least three questions that a jury would have to decide in assessing a plea made on the basis of principles like those of Rawls: (i) is the dissenter's claim that he believed he was justified in what he did a genuinely sincere claim? (ii) is the dissenter's belief that he was justified in what he did reasonable as well as genuinely

sincere? (iii) is this belief correct, in addition to being both reasonable and sincere? Ideally, a defendant would be able to convince a jury that the answer to all three questions was affirmative. However, I think that it will often be impossible to reach agreement on the third question, while the answer to the first two will clearly be yes. Hence I think that proving that one was correct in his conscientious disobedience would be too strong a requirement for acquittal. On the other hand, sincerity in itself seems too weak to qualify as a sufficient condition in and of itself.³³ Hence, establishing that one was sincere in his moral dissent will be a necessary condition for acquittal on my proposal, and establishing that one's sincere dissent was reasonable and/or correct will be a sufficient condition for acquittal.

What does it mean to say that a person's belief with respect to the moral propriety of disobeying the law was a correct and/or a reasonable belief? Suppose, as above, that we accept Rawls's account of when civil disobedience and conscientious refusal are morally justified. We can then say that a sincere belief that one was morally justified in his non-compliance is a correct belief just in case one's belief that the appropriate conditions for noncompliance were fulfilled is correct. Similarly, we can say that a sincere belief that one was morally justified in non-compliance is a reasonable belief just in case one's belief that the appropriate conditions for noncompliance were fulfilled is a reasonable belief. What will count as correct and/or reasonable beliefs about these matters will itself be a matter of controversy, which we shall discuss in a moment. But first a word about my procedure here and the structure of the defense I am proposing.

I am assuming without argument that we can describe at least three kinds of dissent: (i) correct moral dissent (i.e., dissent that is actually justified by correct moral principles); (ii) reasonable moral dissent (i.e., dissent that it is reasonable to undertake, whether it is actually justified or not); and (iii) genuine moral dissent (i.e., dissent that is conscientious or sincere, even if it is neither reasonable nor correct). Obviously, more than one of these descriptions can apply to the same act at the same time.

I am also assuming, for the sake of argument rather than "without" argument, that the standards for when illegal dissent is in fact morally justified are roughly the standards outlined by Rawls in Chapter VI of A Theory of Justice. I happen to believe that Rawls's view is essentially correct, but I cannot defend this view here. The point is that I shall assume that we have these rather rough standards at our disposal. The problem is how we are to apply them, and in particular whether we should allow them to be applied by judges and/or juries in courts of law.

Now since I assume that in every case we shall be concerned with the defendant will have to establish his sincerity or conscientiousness beyond the shadow of a doubt (this is what is meant by saying that sincerity is a necessary condition of a successful legal defense), we are left with four possible kinds of cases, depending on whether the defendant's sincere belief is judged to be correct or incorrect, reasonable or unreasonable. This can be illustrated, with the different cases identified for reference by capital letters, as follows:

Status of a sincere belief that illegal
dissent was morally justified

	Reasonable	Unreasonable
Correct	A	C
Incorrect	B	D

The cases that are of most interest to us are A and B, of course: cases where the defendant's belief was (i) a reasonable and correct belief, and (ii) a reasonable but incorrect belief, respectively. In cases of an unreasonable, incorrect belief, I shall argue that the defense (for acquittal on grounds of morally justified conscientious noncompliance) fails. I am not certain what an unreasonable but correct belief that one was morally justified in disobeying the law would be like (case C), but doubtless we could work out this sort of case if we needed to.

My view is that defenses to the effect that one's illegal dissent falls in either category A or category B ought to be allowed in criminal trials concerned with conscientious disobedience to the law (what Rawls

would call civil disobedience or conscientious refusal), and that if one can make good on the defense (i.e., convince the judge or jury, as the case may be), then one ought to be acquitted outright (in cases of type A) or excused for the offense (in cases of type B). The controversy, of course, is over whether it is reasonable even to allow defenses of this sort.

5. I want to begin assessing the pros and cons of the system I favor by illustrating it with an example of each of the relevant sorts of cases. The clearest case for acquittal on moral grounds (category A) will be where the jury concludes that the defendant acted on sincere moral convictions that were both reasonable and correct.³⁴ Here I should like to say that the individual was fully justified in what he did. An example would be the case of a soldier in Vietnam who refused to obey an order to shoot down what looked like, and turned out to be, a group of peasants.

The clearest case of a failed moral defense (if we exclude for now cases where the defendant cannot even establish his sincerity) would be a case where the jury concludes that the defendant acted on moral convictions (however sincere) that were unreasonable and incorrect (category D). Suppose, for example, that a U.S. citizen had illegally passed important documents to the Axis powers during World War II on the grounds that the Allies were waging a war that threatened to end an era of human progress and peace.

Perhaps the most interesting class of cases are those where the jury cannot agree that the defendant's moral judgment was correct, but where they believe that nonetheless it was reasonable, or at least not unreasonable, under the circumstances (category B). Consider, for example, the case of a young man who refuses to be drafted into the armed forces of any major power like the United States, and to do any form of alternative service, on the grounds that to do so in the context of contemporary international politics is to participate in a militaristic system that is inevitably going to cause more harm than good in the world.³⁵ In cases

like this, where the jury concedes that the view is reasonable even if not clearly correct, we should want them to be empowered to recommend a mitigated or suspended sentence.

I should like to show first that there are no profound difficulties of principle involved in my proposal and, moreover, that far from supporting premiss (8) in the argument of the preceding chapter, or anything like it, Rawls's general theory of politics and society actually supports an arrangement like the one I have suggested. I shall show that if we take seriously the idea of maximizing justice in our social and political institutions, and the idea that an injustice is morally tolerable only if it is required to prevent an even greater injustice, then we shall be inclined to allow defenses like those I have described rather than to prohibit them. In section 7 I shall try to show that the practical difficulties involved in the implementation of a proposal like my own are not severe enough to override its theoretical and moral superiority to systems like our present one, which does not allow moral considerations to be presented as a defense for disobedience to the law, at least to any appreciable extent.³⁶

The principal merit of the system I propose is that it would enable dissenters who were morally justified in their disobedience to the law to state their case and, hopefully, thereby escape the ordinary legal consequences. Of course, such a system might allow too many (or the wrong) people to escape the consequences, and this is a difficulty we shall have to consider. The important point here, however, is that there is a strong prima facie case against punishing justified civil disobedience (see section 3 in Chapter V), and any argument for premiss (8) above and the willingness requirement must show why it can be right to punish in such cases despite this argument. The earlier argument against punishment in cases of this sort was actually based on principles that Rawls himself has articulated -- the principle that an injustice is justified only if it is necessary to prevent an even greater injustice, for example. But there is another line of thought that is relevant here as well.

6. Professor Herbert Packer has argued for a view related to mine by

distinguishing the aims of punishment and the aims of what he calls "the rule of law".³⁷ Suppose we accept the claim that the goal of punishment is the prevention of crime. If we distinguish the aims of punishment (within a legal system) from the ends of the legal system itself (or as a whole), then we can argue that the prevention of crime cannot be pursued without qualification, because there are certain other values that "transcend the goal of crime prevention" as Packer puts it.³⁸ What are these values? According to Packer, in a free society law must ultimately be judged "on the basis of its success in promoting human autonomy and the capacity for human growth and development."³⁹ Not only does "punishment of the morally innocent", as Packer calls it, fail to reinforce one's commitment to the rule of law generally, but a single-minded pursuit of the goal of crime prevention -- excluding moral defenses of the sort sketched above -- "will slight and in the end defeat", Packer argues, "the ultimate goal of law in a free society, which is to liberate rather than to restrain."⁴⁰

Packer's view rests on an emphatic distinction between the aims of punishment and the aims of the rule of law generally, and on the possibility that the two can conflict. Although Packer has a rather special conception of the function of law, it is not an idiosyncratic one. In many ways it is reminiscent of the view adumbrated by H.L.A. Hart in the controversy over the proposal to eliminate the "mental elements" of mens rea and personal responsibility in the law. Hart points out that in the criminal law of nearly all modern states responsibility for certain crimes is excluded or "diminished" by what he refers to as "excusing conditions". On a more general level, this is the doctrine that a "subjective element" or mens rea is required for criminal responsibility.⁴¹ This is related to my own argument in an important way.

Hart is concerned in his article not simply with arguing for the preservation of the concepts of mens rea and personal responsibility but, first, with inquiring into and showing precisely why these concepts are worth preserving -- that is, with analyzing and refuting the so-called "retributive" and "utilitarian" rationalia that have been advanced by

various thinkers in support of the preservation of these concepts -- and secondly, with determining just where their ultimate rationale really lies and of just what sort it really is. Thus, he asks why we value a system of social control that takes mental elements into account and what would be lost if we gave this system up. After a careful analysis of the "retributive" and "utilitarian" attempts to answer these questions, Hart outlines his own answer, which he finally summarizes in the following way:

On this view excusing conditions are accepted as something that may conflict with the social utility of the law's threats; they are regarded as of moral importance because they provide for all individuals alike the satisfactions of a choosing system [a system in which the individual is free to choose to obey or disobey the law and in which he is not accountable for disobedience which he did not freely choose]. Recognition of excusing conditions is therefore seen as a matter of protection of the individual against the claims of society for the highest measure of protection from crime that can be obtained from a system of threats. In this way the criminal law respects the claims of the individual as such, or at least as a choosing being, and distributes its coercive sanctions in a way that reflects this respect for the individual. This surely is very central in the notion of justice and is one, though no doubt only one, among the many strands of principle that I think lie at the root of the preference for legal institutions conditioning liability by reference to excusing conditions.⁴²

It is clear even in this relatively early essay that Hart sees the rationale for the excusing conditions embodied in the doctrine of mens rea and the concept of personal responsibility not in the "retributive" and "utilitarian" justifications strewn throughout the legal and philosophical literature but in something far deeper and much more profound: in the notion of man as a free and "choosing" creature and in the respect for the individual and his freedom to choose that these concepts (of mens rea and personal responsibility) reflect and reinforce.

There is, I think, an important analogy between Hart's arguments for the preservation of the concepts of mens rea and personal responsi-

bility and my own proposal to introduce a legal defense for criminal behavior that is alleged to be morally justified. This analogy is relevant to the critique of premiss (8) in Chapter V as well. Just as Hart claims that there are certain profound "moral gains" that would be lost if we eliminated the principle of mens rea from the criminal law (such as respect for the individual and his freedom to choose, among other things), so there is much to be gained -- that is, society would be "enriched" in certain profound and important ways -- by the introduction of a legal defense for principled or conscientious disobedience to the law.

To begin with, the very same values of respect for the individual and his freedom to choose would be protected. Moreover, we surely want to encourage an individualistic, reflective response to the law in societies like our own, for the alternative is a society that will consist more and more of "moral automatons" who look always for a law or some other socially imposed duty to guide their behavior and who never guide it for themselves and by their own lights. Surely higher and higher levels of moral consciousness enhance the quality of civilization in a given society in obvious ways. Just as a legal system that would permit widespread use of laws of strict liability and eliminate the notions of mens rea and personal responsibility would lose the profound values of respect for the individual, for example, and the others that Hart finds at the bottom of these concepts and of our distaste for such laws, so too, if we simply foreclose the possibility of defenses based on the fundamental principles of political morality, we stand to lose something on which we ordinarily place a great deal of value: respect for individual choice and moral freedom.

A less florid attack on premiss (8) and a defense of the alternative arrangement it precludes could be sketched by trying to answer the following question quite concretely: how are we "enriching" society by allowing moral defenses for criminal conduct? We are doing so, it might be answered, in at least two ways. On one level we are showing a proper respect for moral freedom and autonomy, something that is important and valuable in its own right. But suppose it is contended that we would have to pay too high

a social price for doing this. To this it might well be replied that encouraging a sincere, considered ethical approach to social conduct, even when it may involve crime, is at least arguably in the interests of society itself in the long run. Who can say, for example, what the effect on the national awareness would have been if draft resisters had been allowed to give their reasons, and show evidence for their beliefs, in trials of conscientious crimes related to the awful war in Vietnam? A defense involving a plea of "moral justification" would, in short, be a two-edged sword, protecting the individual, or at any rate, giving him a chance to state his case, and informing the public in the very process, bringing to their attention not only moral convictions but also factual evidence that they otherwise would probably never see.

I would now like to make one last and entirely different general remark with respect to the critique of (8) and in connection with my own proposal for a legal defense of "moral justification". It is a point suggested by Professor Hart in another of his essays in quite another context.⁴³ Hart argues, quite convincingly I think, that despite the fact that our general aim in punishing law-breakers is a utilitarian or "foreward-looking" one, and not a retributive or "backward-looking" one, when we come to actually applying punishment to an individual offender we do not rely on strictly utilitarian considerations. We qualify our initial utilitarian aims, as it were, because justice or fairness seems to demand that whatever the utility of punishing wives and children of law-breakers, for example, in order to ensure greater "deterrence value", considerations of utility shall not prevail at this point and we shall punish only offenders for offenses, thus adopting, in a sense, a quasi-retributive or at least non-utilitarian point of view at this juncture.

Now I think Hart's general point applies to the institution of law itself. Although the general aim of a legal system and its sanctions may be a utilitarian one in some sense (very roughly: the prevention of crime and the preservation of "law and order" in society), certain other moral principles may also make it necessary to qualify that general aim at certain points just as in the case of the institution of punishment.

It may well be that certain forms of strict or extreme "legalism", in other words, while quite acceptable from the strictly utilitarian point of view, may have to be limited or qualified because of other (non-utilitarian) moral principles.

Thus, while there may be more over-all social utility (in some classical sense) in a legal system that allows no departure from the injunctions of the "the law" and the legal obligations it imposes, it can be argued that principles like justice or fairness require us to qualify this system by recognizing and actually introducing a legal defense for principled or conscientious disobedience to the law. And if we fail to observe this limitation or qualification, we may very well be unable to give a morally tolerable account of our institution of law, as Hart puts it.

It seems to me that these remarks are especially apposite in the context of an essentially non-utilitarian theory of political obligation such as that of Rawls. Indeed, Rawls could argue for precluding the qualifications suggested by the preceding remarks, qualifications that undermine premiss (8) in the argument by analogy, only by showing that the willingness requirement results not in a greater balance of "utility" (in the classical sense) but in a greater balance of justice (in his own sense). That is, he would have to show that despite its theoretical merit, a system that allows moral defenses in the courtroom is less effectively just, in the long run, than one that does not. This means that he would have to show that, as a matter of empirical fact, an arrangement that precludes moral defenses is productive of more justice, in the long run, than one that does not; or, at any rate, that among the arrangements available to us, the one that precludes such defenses is likelier to preserve the maximal balance of justice over injustice than an arrangement which allows such defenses.

Rawls does not defend this claim in A Theory of Justice, although I think he needs to support some such contention to sustain his claim that punishing justified civil disobedience can be morally right, or is at least morally tolerable. He is committed, after all, to the principle that an

injustice is morally tolerable only if it is required to avoid an even greater injustice. It is not difficult to imagine the sorts of considerations that could be adduced to support the claim in question, and I wish to speak to them in the following (and final) section.

7. What exactly is supposed to outweigh the apparent injustice of punishing justified civil disobedience? And what makes denying conscientious disobedients a chance to state their case the preferable thing to do, from the point of view of justice? It will be helpful here to remember the analogous argument for requiring compliance with some unjust laws. The argument for the crucial empirical premiss there (see section 8, Chapter IV) was that, up to a point, we had to require obedience even to unjust laws for the sake of the greater balance of justice in the long run. A certain number of unjust laws is inevitable, and it is better (up to a point) to have more or less uniform obedience than to have selective obedience and disobedience. Letting this argument stand for now, what is the analogous argument for premiss (8) in the argument by analogy? Presumably, the claim would be that we have to require a willingness to accept the consequences (and we have to apply them) for the sake of the greater balance of justice. That is, it is inevitable that there will be people who we will agree were justified in disobeying the law, but it's better -- in the long run and from the point of view of maximizing justice -- to disregard pleas based on the moral propriety of disobedience than to entertain them. Why should anyone think that this is so?

To begin with, the alternative system (of hearing pleas of moral justification, etc.) would take too much time. This is the weakest defense of premiss (8), when we remember that we want to maximize justice as well as efficiency, and in any case I shall show below that this difficulty can be obviated in fairly straightforward ways. There are obviously more impressive things to be said on behalf of (8). Thus, it might be alleged that it would be practically impossible to distinguish those who were justified in their disobedience from those who were not. Unfortunately, this defense is not open to Rawls, since presumably the

point of the principles of justice and the general theory of political obligation and civil disobedience that he develops around them is to give us criteria for making exactly these sorts of decisions (criteria, that is, for distinguishing those who are justified in their disobedience from those who are not).

But perhaps Rawls can use a related point to make good on (8). Suppose we do in fact have criteria for meeting the preceding objection. After all, we have Rawls's general theory of justice for institutions and his inchoate but very suggestive "theory of right" for individuals as well as his very detailed theory of the justification of civil disobedience and conscientious refusal. Nonetheless, even if we have these criteria, how can we trust one another to apply them properly in particular cases? They are, after all, merely approximations to the truth on these matters, as Rawls himself repeatedly emphasizes, and only rough approximations at that. They were developed, moreover, with more of a theoretical than a practical intent.

This is not the difficulty, which we have not yet considered, that arises out of the claim that people's conceptions of right and wrong are too divergent for the alternative to (8) to be practicable. We can suppose for the sake of argument that at bottom people do in fact have a common sense of justice. The problem we are considering here is the problem of using that sense of justice, and the particular principles Rawls thinks underlie it, to decide in particular cases who is justified and who is not in his disobedience to the law. We might call this the problem of separating the sheep from the goats, morally speaking. The general problem is clear enough, at least intuitively, but how can it be worked out to provide a compelling defense of premiss (8) in the argument by analogy?

Consider the alternative: allowing a defense such as "moral justification" and basing it, say, on Rawls's general theory of justice and his particular theory of justified noncompliance. Imagine the long, drawn-out disputes we would encounter, in such a system, in the court room and in the jury room.

Notice first that there already are long, drawn-out disputes in the jury room. Would we really be making things any worse -- from the point of view of the jury -- by allowing them to consider moral and political issues? It may be that we would even simplify things for the jury in certain cases. I shall have a great deal more to say about this in a moment; here let me simply say that it is not altogether improbable that a jury's task would be made easier if they were allowed to dismiss a case against a man on the grounds that even if he'd done what he was alleged to have done he was justified in doing it.

It may be argued that the arrangement I favor will be abused and that society cannot take the risk of having unscrupulous and morally perverse persons let off, even at the risk of punishing the righteous and the morally innocent. But this is an odd reversal of a principle that one would have thought was basic to our jurisprudence in this area: better that a thousand guilty men go free than that one innocent man be wrongly punished. Perhaps we want to qualify this platitude. Or perhaps it was meant to apply only to men "innocent of crime" and not men who are admittedly guilty of a (nonviolent) criminal act whom we merely consider to be morally innocent. The latter cannot be the proper move for the advocate of (8), since the moral propriety of punishing such people ("moral innocents") is exactly what is at stake here; surely their moral innocence is relevant, even if it is not conclusive. Perhaps it is relevant, then, to just how much punishment of the morally innocent we are willing to endure in order to safeguard society against conscientious (and convincing) persons with bad moral principles.

This point seems to rely upon postulating an almost overwhelming moral obtuseness in the general public.⁴⁴ It overlooks, moreover, what I have been at great pains to emphasize and shall return to again below: a system that wants to take into account the possibility that a criminal was morally justified in his criminal conduct does not have to begin by letting off all criminals who claim that they acted conscientiously. Indeed, to do so would be madness, and the idea that we have to do so, if we reject premiss (8), rests on the mistaken idea that we either exonerate

conscientious disobedience wholesale or not at all. In fact, the reasonable thing to do would be to give conscientious law-breakers a chance to state their case and then give a thoughtful jury of their peers a chance to evaluate that case.

There may be some danger here, even serious danger, but it is not the danger that an affected jury of satanists, say, would exonerate Charles Manson on grounds that he acted in good conscience. In fact, it is odd that difficulties of this sort should be thought to attach to alternatives to (8) and not to systems just like our present one as well. A jury is in fact, if not in theory, free to find a Charles Manson innocent, if it should happen that a jury is selected all of whose members choose to ignore the defendant's de facto guilt and return a verdict of "not guilty" without even explaining why.

It does seem to follow from my proposal, however, that if a jury composed entirely of racists chose to exonerate an individual who selectively (and illegally) discriminated against Blacks active in the civil rights movement, then it can rightfully do so. There are a number of replies to this difficulty, none of which is conclusive but which will suggest the direction a fully adequate reply would take.

To begin with, any jury that would do this under my system would doubtless do it anyway under our present system. In fact, this sort of thing seems to have happened with a disturbing regularity in the South. What is more important, however, is that we must weigh evils of this sort against the good that would be done in other cases, cases where a dissenter who was morally justified in his disobedience makes a convincing case for his act of disobedience and is not forced to suffer the consequences. Despite the fact that many disobedients who are justified in their actions seem to be quite willing to pay the penalty for doing so, a certain amount of good would be done if they didn't have to, and this good must be weighed against the ill effects of the arrangement I favor.

Another problem with my proposal is the problem of "differential enforcement" which seems to lead to related problems of inequity and uncertainty. Even if we suppose the best about each other (that we share

a common sense of justice, for example, however much our individual views differ in detail), there will be plenty of cases where individual A, say, will be let off by jury Y, while individual B, guilty of the same crime and sincerely using the same moral defense, will not be let off by jury Z, for example.

I am not impressed by the alleged "inequity" of this feature of my system, but I am troubled by what I shall call the problem of "uncertainty".⁴⁵ With respect to the case just described and the unequal treatment A and B have received, I should think that it is better that at least one of the justified disobedients gets off (assuming, of course, that he was justified) rather than neither. It would be best, of course, if judicial decisions were uniform in this respect, and all justified civil disobedience was treated in the same way. But surely the good of the one man getting off outweighs the "inequity" of the other's not getting off, when the alternative is to let neither man off.

I am more concerned, however, about a related but independent problem. It may be that if there were an institutionalized defense for conscientious civil disobedience, people would begin to count on acquittal when they disobeyed the law with the firm conviction that they were justified in doing so. I am not raising here the problem that unscrupulous individuals would tend to abuse this defense (I consider that problem below) but the problem that a man who had relied on the social conscience might be disappointed by the actual outcome of his trial. I'm not sure just how much of a problem this is for my proposal, even though it represents an obvious problem for people contemplating conscientious disobedience. There is no problem, in the system I favor, that an individual who is innocent of even breaking the law might be convicted because a jury didn't like his particular moral and political convictions, since the defense I am proposing could be used only to acquit and not to convict. Under the arrangement I favor, as in our present system, anyone who felt that he had been unjustly convicted, because he wasn't even guilty of committing the crime, questions of possible justification aside, could simply appeal his conviction to a higher court. There is evidence at present

that judges or juries sometimes allow their appraisal of the moral character of the defendant to distort their judgment about the facts of the case.⁴⁶ It does not seem to be that this difficulty would be any more widespread in the kind of system I favor.

Finally, it might be objected that if, as I contend, the American jury in effect already has (and sometimes exercises) the power I want to give it, then it is otiose to do so. This objection faces at least four difficulties. First, this de facto power is one that we ought to certify as a de jure power and not one that we should let slip in occasionally by the judicial "back door". Second, it's not the case that juries uniformly take moral considerations into account as the objection implies, and thirdly, they would be much more likely to do so (and to do so carefully) if they were instructed by the judge that they had a right to do so. Fourthly, and most importantly, by officially certifying "moral justification", for example, as an acceptable legal defense, we would bring the discussion of the moral issues into the courtroom, thus informing the jury of aspects of the case they would otherwise never hear about or, what's worse perhaps, which they would be instructed to disregard if they did hear about them.

In connection with this last point, there is still an apparent problem of time. Speaking practically, ignoring the deeper issues, and bearing in mind the enormous load the courts are already under, is it reasonable to suggest that we complicate and lengthen an already complicated and lengthy process? One obvious reply is that the time factor, even if considerable, does not weigh very heavily against the reasonableness of allowing such a defense, unless it in fact affects the over-all justice of the judicial process. It is irrelevant that the arrangement I favor might be more time consuming than its major rival, unless this feature would tend to make my arrangement productive of less justice, in the long run, than the alternative arrangement favored by most recent thinkers. But perhaps that is just the point; the arrangement I favor would not simply be inefficient, because of the time factor, but would be less effectively just as well (defendants might have to wait even longer than at present, for example, simply to state their case). This

is a serious difficulty. It cannot be conclusively met without specifying exactly how the defense I favor would function in practice and calculating roughly how much time it would add to judicial proceedings, and I cannot make these specifications and calculations here. I would, however, like to make an obvious point. It is not at all clear that the sort of defense I have in mind would significantly alter the amount of time spent in the court room, which is where the problem of time is paramount. The defendant would state his (moral) case, the prosecution could reply, and perhaps the defendant would have a final word. The real complication would presumably occur in the jury room, and hence wouldn't add to the court-room time to any considerable extent. What about time spent in jury room deliberations?

It is important to separate the question of how much more complicated the jury's task would be, under the arrangement I favor, from the question of how much more time consuming it would be. There is no question that in one sense the jury's task would be more complicated under the arrangement I favor; they would be judges, in some cases, not merely of the facts surrounding a particular case of allegedly criminal conduct but of the moral propriety of that conduct as well. And it will often be difficult for them to make a responsible decision in this regard without a great deal of discussion. At some point, however, and I am inclined to think after not too long, it will be fairly clear whether the jury is sufficiently impressed by the defendant's case to acquit him on moral grounds. If they are not, then at least he has had a chance to state his case. If they are, then surely the extra time spent in sorting out cases of justified from unjustified disobedience can be viewed as an intelligent investment in the interests of effecting a greater balance of justice over injustice in a given society.

Two serious practical difficulties remain, difficulties that I have already mentioned and that any arrangement like the one I favor must face. One is the possibility that there is no common core of conviction with respect to the fundamental principles of political morality and hence that it is unreasonable to expect juries to settle cases based on an appeal to

such principles. The easy answer is to say that this is not a problem from the present perspective, since Rawls, at least, assumes that there is, in a society like our own at any rate, a common sense of justice of the kind necessary to meet this objection.

I should like to be able to show, however, not only that Rawls believes this, and hence that he at least ought to accept the arrangement I favor, but also that he is correct in his belief and hence that all reasonable men ought to accept it. I cannot show this here, but I wish to stress -- what is perhaps obvious -- that this is an empirical and not a normative claim. What we need to know, in order to assess the objection at hand, is the extent to which people in a society like our own share certain convictions about the fundamental principles of political morality. Perhaps I am overly optimistic in my own beliefs on this score, but I should like to think that more often than not, when people seem to disagree about these basic principles, it is because they disagree about certain matters of fact -- about the facts of a particular case or the likely consequences of disobedience -- than because they disagree about the principles involved.

A final practical difficulty arises out of the possibility that unscrupulous individuals would abuse a defense like that of "moral necessity", with the result that the balance of justice would suffer rather than being enhanced. Objections of this sort have been considered at several points above, and to a large extent the discussion there is apposite in reply to this particular difficulty. Thus, we must balance the inevitable "utilitarian losses" to which the arrangement I favor would be liable, against the straightforward moral gains it would produce. Similarly, we must ask how high a price we are willing to pay to prevent losses of this sort, and in particular, whether we are willing to pay the price of continuing to disallow in practice a legal defense for conduct that we are willing to agree can be justified in theory.

This, of course, raises again the largest question I have been speaking to throughout the present study: Is it reasonable for the very theo-

rists who have successfully shown that civil disobedience can be morally justified (and hence that some cases of justified civil disobedience can be successfully identified) to insist that the disobedients must accept the consequences of their illegal conduct? I have attempted to show that the best argument for saying that such a position is reasonable is based not on the theoretical merits of the willingness requirement but on the practical difficulties of allowing a legal defense for conscientious disobedience to the law. And I have tried to show that the claim that these difficulties outweigh the apparent injustice of punishing justified disobedience to the law is itself based on empirical presumptions that are dubious at best and that certainly have not been proven.

NOTES -- CHAPTER I

¹For a sample of the literature from a decade ago, when philosophical interest in the justification of civil disobedience was on the rise, see the articles by Hugo Bedau, Stuart M. Brown, Jr., and Richard Wasserstrom in the JOURNAL OF PHILOSOPHY, Vol. 58 (1961). See also many of the essays in Sidney Hook (ed.), LAW AND PHILOSOPHY (New York University Press: New York, 1964), especially the essays by Rawls, Konvitz and Brandt, which are discussed below. Since the early 1960's there has been a profusion of articles propounding the view I am alluding to. For the best recent statement see John Rawls, "The Justification of Civil Disobedience" in Hugo Bedau (ed.), CIVIL DISOBEDIENCE: THEORY AND PRACTICE (Pegasus paperback: New York; 1969) and sections 55-59 of Rawls's book, A THEORY OF JUSTICE (Harvard University Press: Cambridge, 1971). See also Marshall Cohen, "Civil Disobedience in a Constitutional Democracy," THE MASSACHUSETTS REVIEW, Vol. X, No. 2 (Spring, 1969), and "Liberalism and Disobedience," PHILOSOPHY AND PUBLIC AFFAIRS, Vol. 1, No. 3 (Spring, 1971). Carl Cohen has recently tried to tie some of the early literature together in his book, CIVIL DISOBEDIENCE: CONSCIENCE, TACTICS, AND THE LAW (Columbia paperback: New York, 1971), but much of the book is intellectually sloppy and careless; it is worthwhile as a popular gloss of a view defended more carefully elsewhere.

²See the work of Rawls, for example, mentioned in footnote 1, as well as that of both Marshall Cohen and Carl Cohen. Despite differences in approach, and even differences on some substantive points of detail, these authors tend to agree that civil disobedience is justified only when it is necessary to protest a serious breach of justice in society and only when normal legal remedies have been of no avail. It must also, by definition for most recent authors, be a public and nonviolent act, among other things.

³The dissenters include Abe Fortas in CONCERNING DISSENT AND CIVIL DISOBEDIENCE (Signet: New York, 1968) and George Kennan in "Rebels without a Program," reprinted in DEMOCRACY AND THE STUDENT LEFT (Bantam paperback: New York, 1968). Fortas's views are well known; Kennan expresses "a serious doubt", in his essay (originally printed in the NEW YORK TIMES SUNDAY MAGAZINE, January 21, 1968), "whether civil disobedience has any place in a democratic society" (p. 15).

⁴Howard Zinn is an outstanding exception to this generalization -- see his DISOBEDIENCE AND DEMOCRACY (Vintage paperback: New York, 1968), pp. 27-31 -- but Zinn fails to sustain his position with sound arguments. Kai Nielsen has apparently attacked this feature of the received view recently -- see footnote 12, p. 289, in M. Cohen "Liberalism and Disobedience", op. cit. -- but I have not been able to get hold of his paper. Professor Cohen's remarks suggest that Nielsen's arguments are defective.

⁵ See, for example, Carl Cohen, op. cit., pp. 94-101. See also Richard Wasserstrom's well-known essay, "The Obligation to Obey the Law", UCLA LAW REVIEW, Vol. 10, No. 4 (May, 1963) at p. 785 and passim.

⁶ Hugo Bedau, "On Civil Disobedience", op. cit., p. 653. For a more recent statement of Bedau's views see "Civil Disobedience and Personal Responsibility for Injustice", THE MONIST, Vol. 54, No. 4 (Oct. 1970).

⁷ Bedau, "On Civil Disobedience", op. cit., p. 659.

⁸ Ibid., p. 659, footnote 9.

⁹ See Bedau's remarks at p. 661 of his early article and at p. 519 of the later article. See also the definition of Rawls and M. Cohen in the essays mentioned in footnote 1 above.

¹⁰ See the definition in the early article mentioned above. In the later article I think Bedau pretty clearly makes the willingness requirement an explicit part of his definition or "characterization" of civil disobedience; see p. 519 of that article.

¹¹ I am not suggesting that Bedau in fact tries to do this; I merely wish to emphasize that it would be a mistake for anyone to try to defend the willingness requirement in the way described.

¹² Marshall Cohen, "Civil Disobedience in a Constitutional Democracy," op. cit., p. 213.

¹³ Ibid., p. 214.

¹⁴ Ibid., pp. 214-215.

¹⁵ The position I shall be concerned with rebutting, then, is that which makes acceptance of the legal consequences a necessary part of the moral justification of the illegal act and not merely a part of its definition or its reasonableness from a tactical point of view. It will not always be easy to keep these components separate, and, indeed, it may not always be necessary or reasonable to try to do so with absolute rigor. But to the extent that I can do so, I shall try to isolate each thinker's reasons for making a willingness to pay the penalty a moral requirement in the justification of civil disobedience as opposed to a merely tactical or definitional one.

¹⁶ Stuart M. Brown, Jr., op. cit., p. 676.

¹⁷ Ibid., p. 678.

¹⁸ Ibid., pp. 678-679.

¹⁹ See Marshall Cohen, e.g., "Liberalism and Disobedience," op. cit., p. 298.

²⁰In Chapter II I consider the claim that there is some sort of conceptual (or other) connection between the requirement that civil disobedience be nonviolent and the willingness requirement. Brown does not attempt to connect these notions, however, so I do not consider the appropriate arguments here.

²¹Brown, op. cit., p. 679.

²²I return to this point, and consider it at much greater length, in Chapter II, where I consider a related claim of John Rawls's, and in Chapter III, where I consider a claim recently advanced by Richard Wasserstrom and also by Marshall Cohen. Here I simply want to meet Brown's explicit arguments. There I shall be somewhat more generous, trying to see if anything more can be said for this claim, even if Brown doesn't say it.

²³Brown, op. cit., pp. 679-680.

²⁴Ibid., p. 678.

²⁵In this regard, see his remark quoted earlier: "a public protest in which the participants injure others, destroy property, and resist arrest is no less an act of civil disobedience; it is merely an unjustified one."

²⁶Milton Konvitz, "Civil Disobedience and the Duty of Fair Play", in Sidney Hook (ed.), LAW AND PHILOSOPHY, op. cit., p. 26.

²⁷Ibid., p. 27.

²⁸Ibid., p. 28.

²⁹I return to Rawls's paper in Chapter IV below.

³⁰Konvitz, op. cit., pp. 26-27.

³¹Ibid., p. 27.

³²Marshall Cohen, "Liberalism and Disobedience," op. cit., p. 286.

³³Carl Cohen, op. cit., pp. 76-91.

³⁴Cohen favors a broad but not very clear form of consequentialism (perhaps what Moore calls "agapistic utilitarianism") with respect to the justification of civil disobedience: civil disobedience is justified when it produces more good than evil for society (where the evil of disobeying the law is part of our calculation). Cohen recognizes that it will be difficult, in practice, to determine when this is likely to be the case, but he apparently believes it can be the case from time to time. As I shall show in Chapter III, Cohen, like most recent theorists, is quite

conservative with respect to the basic issues here: the potential disobedient bears an enormous burden of proof in justifying his conduct; nonetheless, such conduct can be justified, according to Cohen.

35 I am grateful to Barbara Onutz for help on this and other points connected with Cohen's views on the willingness requirement.

36 I am indebted to Larry Peterson for helping me to get clear on this point.

37 Cohen writes as though a retributively-minded judge could consistently acknowledge that the disobedient was morally right in what he did while insisting that he was especially liable to punishment since his violation of the law was defiant and deliberate (see pp. 80-81, especially). Professor Marshall Cohen has countered, in another context, that "it is a particularly barbarous fallacy to suppose that the government owes the disobedient his just portion of punishment" ("Liberalism and Disobedience", op. cit., p. 288). As I try to show in Chapters V and VI below, I think the view that civil disobedience, even when it is justified, justly merits punishment, is intimately connected with the view that we cannot allow conscientious disobedients to state their moral case in the courtroom.

38 It is important to avoid ambiguity here. My claim is not merely that the disobedient has acted rightly in that he has acted conscientiously and with confidence that he is right in what he's done. I am assuming, as Cohen does, that there will be cases where the disobedient believes, and believes correctly, that he has acted rightly in disobeying the law. The question is whether it makes sense to exact retribution in cases of this sort. Unfortunately, this does not take into account cases where the dissenter acts in good faith but is wrong in his beliefs about the moral propriety of his disobedience. Whether retribution would be justifiable in cases of this sort is a question I cannot answer here. It is also a question I need not answer, however, in order to refute Cohen's simpler point.

NOTES -- CHAPTER II

¹Marshall Cohen, "Liberalism and Disobedience", PHILOSOPHY AND PUBLIC AFFAIRS, Vol. 1, No. 3 (Spring, 1972), p. 285.

²Ibid., p. 286.

³Ibid.; see the text at pp. 287-288.

⁴Ibid., p. 293.

⁵Ibid., pp. 293-294.

⁶Ibid., p. 297.

⁷See the text at p. 294, for example.

⁸Page-references in parentheses are to Rawls's book, A THEORY OF JUSTICE (Harvard University Press: Cambridge, 1971).

⁹See the text at p. 363, for example.

¹⁰See the text at p. 365 in this connection.

¹¹Rawls does not use this particular locution, which is something of a barbarism, but it captures the thrust of his thought and will enable me to make my subsequent points rather simply.

¹²In fact, once one gets down to correlating points on the reaction spectrum with specific points on the circumstances continuum, and explaining exactly why one correlation is more appropriate than another, the enterprise is not nearly so simple. This is why theories of civil disobedience, much less theories of "direct action" or even of revolution, can be so controversial. How bad must things be on the circumstances continuum to justify a given form of dissent on the reaction spectrum? Rawls himself is explicitly concerned, in his book, with only two forms of dissent: civil disobedience and "conscientious refusal".

¹³This does not necessarily mean that circumstances have to change materially; circumstances may remain the same, but after a certain length of time, we've moved on the circumstances continuum and circumstances have changed in that sense.

¹⁴Of course, B will not resist the consequences by means of violence or secrecy, since this would obviously violate the requirements of nonviolence and publicity. I return to this point below.

¹⁵ Rawls's explicit treatment of the willingness requirement in his book is based on the claim that to be justified civil disobedience must not only be public and nonviolent but that it must also be disobedience "within the limits of fidelity to law". In asking, below, what is the connection between these requirements for justified civil disobedience and the requirement that the disobedient be willing to accept the full legal consequences of his act, I shall be dealing with what I call Rawls's "explicit" argument for the willingness requirement (it is the argument that occurs explicitly in this context in his book). In Chapter V I shall show that in fact there is quite a different argument (for the willingness requirement) implicit in Rawls's book and some of his earlier writings. I shall call this "implicit" argument for the willingness requirement "the argument by analogy", since it proceeds by analogy to Rawls's argument for the duty to comply with (some) unjust laws (section 53), which I consider in Chapter IV.

¹⁶ As I have already shown, Rawls says explicitly that such alternatives can be justified but indicates that his concern is with situations where such extreme forms of dissent would not be justified but where civil disobedience would.

¹⁷ See especially pp. 366-367. Civil disobedience is in fact "at the boundary of fidelity to law", according to Rawls; that is, it is the last step short of measures which repudiate fidelity to law.

¹⁸ A society is "reasonably just", on Rawls's view, when it is as just as it is reasonable to expect under the circumstances. Given Rawls's general theory of justice, this means that a society is reasonably just when it comes as close to satisfying the two principles of justice as it is reasonable to expect under the circumstances of normal political life. I return to this point and discuss it in considerable detail in Chapter IV below.

¹⁹ I'm sure that Rawls would consider the situation that American blacks found themselves in during the 1960s an example of the sort of situation described in the text. There were certain laws that blacks were obviously justified in protesting, even illegally. But I doubt that Rawls would say that American society was so unjust in that period as to forfeit any claim whatsoever to legitimate moral authority.

²⁰ Notice that it is not always easy to say just where a given society stands on the so-called "circumstances continuum" at a given point in time. In our own country, for example, so-called "radicals" often differ from less extreme political activists in exactly this respect: the former believe, while the latter do not, that American society is more or less totally unjust. Hence they are willing to resort to violence and revolutionary tactics to achieve their ends. And they believe that the state of American society justifies their doing so.

²¹ See section 38 of his book, "The Rule of Law", where Rawls explains

this notion. This is an important concept in Rawls's conception of a well-ordered society, and I try to work it out in some detail in Chapter IV below.

²²I am indebted to Michael Bratman for clarification on this point.

²³I shall have a great deal more to say about these matters in Chapter IV. For the present I wish to accept the claim that we have this "natural duty" to support and foster just institutions.

²⁴See his essay, "Legal Obligation and the Duty of Fair Play" in Sidney Hook (ed.), LAW AND PHILOSOPHY (New York University Press: New York, 1964).

²⁵In Chapter V, however, I return to this point and consider the possibility that there is another, more profound analogy here and, hence, a far more impressive argument for the willingness requirement.

²⁶See Charles Fried, "Moral Causation", HARVARD LAW REVIEW, Vol. 77 (1963-1964).

²⁷Ibid., pp. 1268-1269.

²⁸Ibid., p. 1269.

²⁹Ibid.

³⁰Ibid.

NOTES -- CHAPTER III

¹It is this tradition that I had in mind above (see Chapter I, Sect. 2) when I remarked that if someone wishes to argue that society is right in punishing me, and that I have an obligation to accept that punishment willingly, when I have rightly refused to obey an unjust law, it would seem that the proponent of this view ought to bear the burden of proof. In particular, we should like to know why it is right to punish right conduct, even if the latter is illegal conduct. This burden of proof, however, is seldom discharged in recent theorizing about civil disobedience; Rawls's theory, as we shall see, is an exception.

²See Bedau's review of Carl Cohen's book (op. cit.) in THE JOURNAL OF PHILOSOPHY, Vol. 69, No. 7 (April 6, 1972) at pp. 184-186. "Cohen's book," writes Bedau, "like the orthodox theory of civil disobedience it expounds and like the paradigm practice of civil disobedience we owe to Gandhi and King, never calls this tradition into question." (Bedau is alluding here to what he has just called "the Liberal-Conservative tradition of Respect for Lawful Secular Authority".) "Yet just see how everything totters," he continues, "if only we voice our doubts: why must the individual contemplating disobedience begin with any burden of justification whatsoever except in a court of law (no doubt the unavowed model in terms of which the "philosophical problem of justification" is developed)? Why shouldn't the initial justificatory burden always be borne by those who would oppose, arrest, try, convict, sentence, and punish the conscientious disobedient?" (p. 186; Bedau's emphasis).

Professor Joel Feinberg is mistaken, I think, when he implies that Bedau is merely calling into question here "the discredited view" that there is a moral presumption in favor of obedience to any law, however instituted and enforced, whatever its provisions, and so forth (see Feinberg's review of Rawls's book in the JOURNAL OF PHILOSOPHY, Vol. 70, No. 9 at p. 269). Certainly Bedau is questioning this (justly) discredited version of extreme "legalism". But as the last two sentences in the quotation above make clear, Bedau is also asking why the individual should be forced to assume any burden of (moral) proof in his disobedience. This question cuts more deeply than Feinberg's formulation, which is also a question Bedau is raising, to be sure. The principal merit of Rawls's theory of political obligation, as I shall show, is that he provides an answer to the question of why there is even a moral presumption in favor of obedience to the law, in more or less "just" democratic societies at any rate.

³John Rawls is an example. In an early article on "Legal Obligation and the Duty of Fair Play" (op. cit.), Rawls writes that "I shall assume, as requiring no argument, that there is, at least in a society such as ours, a moral obligation to obey the law" (op. cit., p. 3); and

in his recent book Rawls claims that "the injustice of a law is not, in general, a sufficient reason for not adhering to it any more than the legal validity of legislation...is a sufficient reason for going along with it" (pp. 350-351).

⁴See Carl Cohen, CIVIL DISOBEDIENCE (op. cit.).

⁵See the articles in the (1961) APA symposium on civil disobedience mentioned in Chapter I as well as the articles in the (1963) NYU symposium on law and philosophy reprinted in Hook, op. cit.

⁶Richard Wasserstrom, "The Obligation to Obey the Law", op. cit., pp. 782-783; Wasserstrom's emphasis.

⁷I am extremely grateful to James Jeffers for forcing me to get clear on this and many subsequent points connected with Wasserstrom's enterprise in "The Obligation to Obey the Law".

⁸See "The Obligation to Obey the Law" at p. 781. This view of traditional democratic political theory is shared by Professor Marshall Cohen; see the first paragraph of his "Civil Disobedience in a Constitutional Democracy", op. cit., and "Liberalism and Disobedience", op. cit., at p. 283.

⁹See p. 781, footnote 2, in Wasserstrom's essay. Wasserstrom attributes the view in question to Austin, Hume, and Locke.

¹⁰A TREATISE OF HUMAN NATURE, Book III, Sects. 9 and 10.

¹¹Speaking of possible exceptions to the general obligation to obey the law, Hume writes:

Our general knowledge of human nature, our observation of the past history of mankind, our experience of present times; all these causes must induce us to open the door to exceptions, and must make us conclude, that we may resist the more violent effects of supreme power without any crime or injustice (Bk. III, Sect. 9).

"'Tis certain," Hume writes later, "that in all our notions of morals we never entertain such an absurdity as that of passive obedience, but make allowances for resistance in more flagrant instances of tyranny and oppression" (*ibid.*). Hume does not say just how "flagrant" these instances of tyranny and oppression must be to justify disobedience and just what sorts of disobedience would be justified under various conditions. But see my remarks below.

¹²See, for example, Richard Brandt, "Utility and the Obligation to Obey the Law" in Hook, op. cit.

¹³This point was illustrated rather nicely by two papers at the N.Y.U. Institute on "Law and Philosophy" in 1963 (reprinted in Hook, op. cit.). Two prominent contemporary philosophers, Richard Brandt and John Rawls, presented systematic analyses of the nature of our obligation to obey the law. Interestingly, their accounts were roughly utilitarian and non-utilitarian, respectively. Rawls indicates explicitly that his concern is "solely...with the grounds for our moral obligation to obey the law," and his central thesis "...is that the obligation to obey the law is a special case of the prima facie duty of fair play." The pertinent point for present purposes can be gleaned from two remarks Rawls makes. At the very beginning of his paper Rawls says: "I shall assume, as requiring no argument, that there is, at least in a society such as ours, a moral obligation to obey the law, although it may, of course, be overridden by other more stringent obligations." Later he writes: "...the principles of justice are absolute with respect to the principle of utility...[i.e.] our obligation to obey the law...cannot be overridden by an appeal to utility, though it may be overridden by another duty of justice." Thus, the same moral principle whereby, according to Rawls, we incur an obligation or duty to obey the law can always be imagined to make it morally justifiable to disobey the law in certain circumstances.

Professor Brandt's account of the nature of our obligation to obey the law is analogous to Rawls's account in the relevant way, although the moral principles on which Brandt would base that obligation are quite different from Rawls's. "I shall argue that utility of consequences is of more importance than Rawls appears to think," Brandt writes. In fact, for Brandt utility is at the heart of our obligation to obey the law: "...if a law has a function in maintaining the social utility at the highest possible level, there is a prima facie obligation to obey that law." But of course this obligation is not absolute, since "it does not follow that reflection will sustain the utility of prima facie obligation to perform." The circle, then, is complete. The obligation to obey the law is based, for Brandt, on considerations of social utility; the justification for disobeying the law is based on precisely the same principle.

¹⁴See Wasserstrom, op. cit., p. 780.

¹⁵Perhaps an autobiographical remark would not be out of order here. My original motive in undertaking the present study was to evaluate a variety of arguments for what I call the "willingness requirement" and to show that they are simply not sound. This is still my principal aim, but in the course of my research I have discovered that behind the willingness requirement lies a more general -- and very common -- conception of the relation between citizen and state, at least in societies that one might call "reasonably just". It is the most extreme form of this view -- that disobedience to the law is never justified in such a society -- that Wasserstrom considers in "The Obligation to Obey the Law". This version of "legalism", however, is pretty clearly false. As we shall see, though, modified and much more sophisticated versions are widely received among contemporary political theorists, in our own

country at any rate. I shall eventually show that this phenomenon -- "legalism", as I call it -- is of much more than historical and academic interest. It accounts, in large part, for the extreme conservatism of liberal writing on civil disobedience, and, in particular, provides the most plausible defense of the willingness requirement that is available to proponents of the received view. In Chapter V I construct an argument for the willingness requirement that, as far as I know, none of its proponents has ever actually advanced but which represents the best defense available to them for their doctrine. This argument is developed from recent work of Rawls; hence, the discussion in this chapter and the next is pertinent both to the question of the rationale of restricted legalism with respect to legislative decisions as well as to the formulation of the argument for extreme legalism with respect to judicial decisions in Chapter V.

¹⁶Subsequent to writing this material I read Marshall Cohen's essay "Liberalism and Disobedience", in which he too points out that "the Socrates of the APOLOGY and CRITO is not standing trial as a civil disobedient" (p. 289). I disagree with Cohen, however, when he suggests that since Socrates has really not broken the law, his willingness to accept the punishment cannot be understood as it should be in cases of civil disobedience. To be sure, the cases are different: Socrates has presumably not broken the law; the civil disobedient has, ex hypothesi, done so. But I shall try to show that Socrates' arguments for accepting the consequences of a fair trial in a just society, even when one is innocent of any illegal conduct, bear directly on the claim that one must accept the consequences of justifiable civil disobedience. Cohen and I may not actually disagree here since, as I have shown in Chapter II, his conception of the role of the willingness requirement in theories of civil disobedience is not the one I am opposed to. What's more, Cohen agrees that "although Socrates is not on trial for having committed a civilly disobedient act, it is true that some of his arguments if successful, would bear on the predicament of the civil disobedient and are therefore relevant to our theme" (p. 290).

¹⁷Plato, EUTHYPHRO, APOLOGY, and CRITO, trans. F. J. Church and R. D. Cumming (Library of Liberal Arts: Indianapolis and New York; 1956), p. 57.

¹⁸Ibid., p. 58.

¹⁹Ibid., p. 56.

²⁰Ibid., p. 57.

²¹Ibid., p. 65.

²²Ibid., pp. 58-59.

²³Ibid., p. 58.

²⁴Ibid., p. 60.

²⁵Ibid.

²⁶Ibid.

²⁷Ibid.

²⁸Ibid.

NOTES -- CHAPTER IV

¹See A THEORY OF JUSTICE, op. cit., Chapter VI, especially section 53.

²I shall assume, following Rawls here for the sake of argument, that we can make sense of the distinction between so-called prima facie duties (or "duties other things equal", as Rawls sometimes calls them) and "actual" or "on-balance" duties ("duties all things considered", as he calls them). I shall assume, with Rawls and a few other recent writers, that there is an important distinction between moral obligations, properly so called, and moral duties. On the first distinction see A THEORY OF JUSTICE at pp. 340-341. For the second point see pp. 113 ff., ibid., and the articles mentioned in footnote 29, p. 113.

³See A THEORY OF JUSTICE at pp. 115-116 and p. 335 ff.

⁴Ibid., p. 335.

⁵Thus, in a full analysis we should want to ask whether we actually have a natural duty to support just institutions, for example, and whether (if we do) Rawls's account of this matter is correct (see section 51 of his book). We would also need to get clear on the status of premiss (2): is this a stipulative claim or a substantive claim of some sort, for example? We would surely have to determine whether (4) is correct, as a plain matter of empirical fact. These are just a few of the questions I am passing over here.

⁶The only alternative is to extract this notion from premiss (4). But this would make (4) either plainly false, if the new claim were that absolute compliance is required, or intolerably vague, if we try to gloss this distinction entirely.

⁷See footnote 2 above on the notion of prima facie duties.

⁸See A THEORY OF JUSTICE, sections 18 and 52; see also Rawls's early article, "Legal Obligation and the Duty of Fair Play", op. cit.

⁹See A THEORY OF JUSTICE at p. 350.

¹⁰Ibid.

¹¹See Rawls, "Legal Obligation...", op. cit., and Charles Fried, "Moral Causation", op. cit.

¹²Rawls himself implicitly (but perhaps unwittingly) acknowledges

this when he remarks in a footnote that the account of the moral basis of civil authority in his book differs from that in his early article ("Legal Obligation...") in that the principle of fairness plays only "a secondary role" in the book, where "the natural duty of justice is the main principle of political duty for citizens generally" (footnote 13, p. 353).

¹³A THEORY OF JUSTICE, p. 4.

¹⁴Ibid., p. 355.

¹⁵Ibid.

¹⁶These conditions are not to be confused with the features of civil disobedience which are part of its definition, according to Rawls -- e.g., the fact that it must be a public and nonviolent act. See A THEORY OF JUSTICE at p. 364. The necessary conditions for justified civil disobedience are sketched in section 57, pp. 371-377.

¹⁷Ibid., section 59.

¹⁸Ibid., p. 380.

¹⁹It is important to notice that they would still be legally bound to those results and, for all we've said so far, to the appropriate legal sanctions. But see Chapter VI below.

²⁰Unjust from the point of view of the first principle of justice, that is, or the second part of the second.

²¹See A THEORY OF JUSTICE, section 38, "The Rule of Law". On the alleged "instability" of the system I favor see section 53 and also p. 367.

²²Of course the latter point might eventually make us want to change our judicial system so that such disobedients could make their case in court. I suggest such changes in Chapter VI below.

NOTES -- CHAPTER V

¹I have not emphasized the degree to which this assumption pervades recent liberal political theory. Thus, Marshall Cohen writes: "It must not be supposed...that the civil disobedient's position implies that he will never submit to the requirements of an unjust law. In fact, the citizen in a democracy often has a moral obligation to do precisely that" ("Liberalism and Disobedience", *op. cit.*, p. 288). Without explicitly saying that the prima facie obligation to obey the law can extend as far as an actual obligation to obey unjust laws, Joel Feinberg makes it clear that he too shares the view that there is a moral presumption in favor of obeying the law in reasonably just societies: "...there is normally a presumption against disobeying the law in a just, or near-just, society -- not an unconditional moral prohibition, but a kind of standing case that must be overridden in a given instance by sufficient reasons" (see his review of Rawls's book, cited above, p. 269).

²Professor Bedau is speaking to this same point, I think, albeit rather obliquely, when he remarks: "Cohen has shown that it is possible to write a brief and intelligent book on civil disobedience, provided one leaves intact Liberal-Conservative assumptions about our 'prima facie duty to obey...[the] law.' No one else should trouble to do this again. Some may wonder why philosophers do it at all" (see his review of Carl Cohen's book, mentioned above, p. 186).

³The final remark ("in the kinds of cases we're concerned with") is crucial, of course. We're presumably at a point on the "circumstances continuum" that is beyond the limits of tolerable injustice but not so far beyond those limits as to justify anything more (on the "reaction spectrum") than civil disobedience as Rawls understands it. This assumption is at the bottom of the discussion throughout this chapter. Thus, when I say that Rawls insists on a willingness to accept the consequences, even when one is justified in one's disobedience, I mean when one is justified in civil disobedience but nothing more. Rawls himself acknowledges that at some point (justified rebellion being the clearest case) it would be absurd (and a moral error) to characterize disobedience as justified and yet to insist on a willingness to pay the penalty.

⁴For simplicity's sake, I have omitted a critical distinction -- the distinction between so-called "direct" and "indirect" disobedience. Thus, I sometimes write as though whenever breaking a particular law is justified it is because that law is itself beyond the limits of tolerable injustice. Strictly speaking, this is only true in cases of "direct" disobedience. In cases of "indirect" disobedience we are indeed confronted with intolerable injustice, but that injustice does not reside in the law that is broken. My remarks below are intended to apply to both

sorts of cases, even when, for the sake of simplicity, I write as though it is the law actually disobeyed that is beyond the limits of tolerable injustice.

⁵The following remarks are based on a point originally suggested to me by David Schweickart.

⁶Again, we are assuming that the limits of tolerable injustice have been passed, but only just passed. That is, nothing more than civil disobedience is appropriate, else it would not even be arguable that the disobedient ought to pay the penalty. (It should also be recalled that I intend my remarks to apply, mutatis mutandis, to cases where the offensive law is not directly the object of disobedience.)

⁷Rawls apparently believes that they should, and would, but of course he does not explicitly say so, since the argument by analogy is my own invention. I believe that it is clearly in the spirit of the Rawlsian enterprise, however, and I believe that it affords Rawls a far more reasonable defense for the willingness requirement than any of the explicit arguments considered in Chapter II above.

⁸See footnote 3 above for the explanation of this parenthetical remark.

⁹The argument that justification entails nonpunishment is designed to bring this out more perspicuously, of course.

¹⁰As we shall see below, what this amounts to in practice is the claim that a greater balance of justice over injustice is achieved, in the long run, by excluding legal defenses based on appeal to moral principles -- even the so-called principles of political morality -- than by allowing such defenses in the courtroom.

¹¹A THEORY OF JUSTICE, op. cit., p. 367. Rawls is not talking explicitly about what I call the "willingness requirement" here, but there is clearly no distortion in applying his remarks to that problem.

¹²Professor Joel Feinberg has remarked to me privately that Rawls may be willing to give up the willingness requirement without a fight. This would be quite congenial to me, although it would make the present study rather academic (to put it mildly) in parts.

¹³Rawls softens his stance at one point by saying: "Courts should take into account the civilly disobedient nature of the protester's act, and the fact that it is justifiable (or may seem so) by the political principles underlying the constitution, and on these grounds reduce and in some cases suspend the legal sanction" (A THEORY OF JUSTICE, p. 387). Rawls then mentions Ronald Dworkin's well-known essay, "On Not Prosecuting Civil Disobedience", in a footnote. I reply to Dworkin, and implicitly to Rawls, in the next chapter.

¹⁴Of course, the phrase "the principle of institutional settlement" might well be thought of as a generic term for institutionalized decision-procedures, in which case it would include the principle of majority rule. I use the phrase much more narrowly here, to serve as the judicial analogue of the principle of majority rule.

¹⁵I say "moral and political defenses", but I am actually interested in defending only the possibility of allowing "moral" defenses, strictly speaking. Though this is not the place to do so, I think it could be shown that one of the principal advantages of Rawls's theory of justice is that he explains the relationship between so-called individual ethics and political morality. Thus, he speaks of civil disobedience as a "political" act, but, of course, it is also a moral act. Professor Marshall Cohen has written elsewhere of "the principles of political morality". The point is that the principles underlying political action are moral principles, as Rawls's theory demonstrates. To underline this point, I shall often speak of "moral and political defenses" for disobedience to the law, where I mean defenses that appeal to the ultimate principles of political morality. Rawls's principles of justice are an example of what I have in mind here.

NOTES -- CHAPTER VI

¹Richard Wasserstrom, THE JUDICIAL DECISION (Stanford University Press: Stanford, 1961), p. 1.

²I explain what I have in mind here in more detail in section 3 below, where I elaborate Rawls's theory of justified noncompliance (his theory of civil disobedience as well as his theory of conscientious refusal) and use it to explicate the notion of injustice used here.

³I have deliberately refrained from saying that his beliefs must be "correct", since this would be too rigorous (and unreasonable) a requirement in cases of the sort I have in mind. On the other hand, we need a stronger requirement than sincerity alone, for obvious reasons. Hence I argue that in order to be acquitted the defendant must convince the jury that his moral convictions are reasonable as well as genuine or sincere. See section 3 below for the details of my proposal.

⁴See almost any recent article on disobedience to the law in both the popular and professional journals. See in particular the articles discussed in Chapters I and II above.

⁵My interest here is partly personal. I was present when a federal judge instructed my unfortunate friend, David Malament, that his reasons for refusing induction were irrelevant to his guilt or innocence before the law. Mr. Malament was advised to speak only to the question of whether he had in fact refused induction and the jury was instructed to disregard other remarks. Since Mr. Malament was endeavoring to explain why he had in fact refused induction, he was in effect pleading guilty.

⁶Such considerations might of course be taken into account in sentencing, or they might not. I discuss this point in detail in section 2 below.

⁷Notice that by using Rawls's theory of noncompliance to explain what would count as "good reasons", I obviate certain serious problems that arise for proposals like my own. In particular, I severely limit the kinds of cases in which a moral defense could be raised in a court of law as well as the kinds of considerations that are appropriately raised there.

⁸Ronald Dworkin, "On Not Prosecuting Civil Disobedience", THE NEW YORK REVIEW OF BOOKS, Vol. 10, June 6, 1968. This article has recently been reprinted in Jeffrie Murphy (ed.), CIVIL DISOBEDIENCE AND VIOLENCE (Wadsworth: Belmont, Calif., 1971). Page references below are to the NEW YORK REVIEW OF BOOKS.

⁹Ibid., p. 14.

¹⁰Ibid.

¹¹The appeal to utility is implicit in the allusion to the "loss" our society suffers if it punishes in these cases; I think Dworkin has another principle in mind -- perhaps a principle of "fittingness" or equity -- when he says that an obvious reason for not punishing in such cases is that the resisters "act out of better motives", etc.

¹²Obviously, Dworkin did not have this point of view in mind in making his proposal, and my remarks should not be misunderstood as a criticism of Dworkin's efforts to propose a workable and realistic solution to an urgent moral and political problem at a time when even men of good will were polarized into bitterly opposed factions. Dworkin was not writing about legal change or reform; as I understand it, his thesis was that to some extent our legal system already had means for resolving the problem of what to do when some of our most conscientious and loyal citizens found it necessary to act in open defiance of the law.

¹³But see Dworkin's addendum to his article in Murphy, op. cit., where he acknowledges that far from taking his advice, the Nixon administration actually increased the number of prosecutions for draft offenses. I saw this postscript to his article only when this manuscript was in the typist's hands.

¹⁴In this connection, however, see footnote 12 above. Dworkin was obviously concerned to make a proposal that would involve a minimum of controversy and a maximum of justice by suggesting a way that the government could respect the resister's views without acknowledging that they were correct.

¹⁵Certainly nothing said thus far in behalf of the willingness requirement is a very compelling reason for not giving them this chance. The central question in the latter half of this chapter will be whether doing so would involve us in a legal arrangement that is obviously less effectively just than an arrangement that prohibits moral defenses for otherwise illegal conduct.

¹⁶I am not suggesting that Dworkin himself holds this view. It is convenient, however, to mention it here, since it faces problems similar to the ones his view faces and also because it is a very popular view.

¹⁷See, for example, the views of former Supreme Court Justice Abe Fortas in CONCERNING DISSENT AND CIVIL DISOBEDIENCE, op. cit.

¹⁸Jon M. Van Dyke, "The Jury As a Political Institution", THE CENTER MAGAZINE, Vol. III, No. 2.

¹⁹Ibid., p. 11.

²⁰Ibid., p. 17.

²¹Ibid.

²²Ibid., p. 18.

²³Ibid., p. 18.

²⁴Ibid.

²⁵Ibid., pp. 19-20.

²⁶Ibid., p. 19.

²⁷Ibid. Van Dyke even finds an example of a system within the United States where, as he says, "the jury is given the power of mercy but not of vengeance." This is in Maryland, where the Constitution reads: "In the trial of all criminal cases the jury shall be the judge of the law as well as the facts." Under this provision, Van Dyke writes, the jury is given the following instructions in every criminal case:

Members of the jury, this is a criminal case and under the Constitution and the laws of the State of Maryland in a criminal case the jury are judges of the law as well as of the facts in the case. So that whatever I tell you about the law, while it is intended to be helpful to you in reaching a just and proper verdict in the case, it is not binding upon you as members of the jury and you may accept the law as you apprehend it to be the case.

²⁸Ibid., p. 26.

²⁹Ibid.

³⁰Ibid.

³¹See his "Liberalism and Disobedience", for example, op. cit., p. 298.

³²Of course, showing that a particular law or policy violated one or the other of these two principles would not be sufficient to make good a defense of "moral justification". For the defendant would have to convince the jury that the breach of these principles was indeed a serious one, that he had acted in defiance of the law only after he had failed to win his rights through legal channels, that his illegal acts were appropriate to the situation and not wantonly criminal, and so forth. In short, the disobedient, or his attorneys, would have to convince the jury that the act in question, while manifestly illegal, was morally justifiable in light of theories like that of Rawls, Cohen, and

others. I take up this problem again below.

³³For a discussion of sincerity alone as a sufficient condition for acquittal see Hugo Bedau, "On Civil Disobedience", THE JOURNAL OF PHILOSOPHY, Vol. 58, No. 21 (October 12, 1961), p. 655. Suppose a jury decides that an instance of moral dissent was motivated out of a genuine concern with the possible injustice of a law or policy, but that that dissent was neither correct nor reasonable. Although we may want to allow the defendant's sincerity to count as a mitigating factor, I do not think that sincerity alone should be thought of as either an excusing or a justifying condition before the law. My position is a conservative one in this respect.

³⁴A slightly weaker case would be one in which the jury concludes that the defendant's moral convictions were reasonable and at least not clearly incorrect.

³⁵This sort of case is suggested by Rawls. See A THEORY OF JUSTICE, op. cit., pp. 381-382.

³⁶I shall not be able to establish this second claim (about the practicability of the arrangement I favor) as firmly as I should like, since ultimately it depends upon empirical data that are not yet available. However, I shall show that the burden of proof is on proponents of alternatives to the arrangement I favor, that they have not discharged this burden of proof, and that the likelihood is that they cannot do so, since empirical investigation is more likely than not to vindicate the practicality of the arrangement I have suggested.

³⁷Herbert Packer, THE LIMITS OF THE CRIMINAL SANCTION (Stanford University Press: Stanford, 1968), p. 65.

³⁸Ibid., p. 63.

³⁹Ibid., p. 65.

⁴⁰Ibid., p. 66.

⁴¹H. L. A. Hart, PUNISHMENT AND RESPONSIBILITY (Oxford University Press: New York, 1968), p. 31.

⁴²Ibid., p. 49.

⁴³See his "Prolegomenon to the Principles of Punishment" in PUNISHMENT AND RESPONSIBILITY, op. cit.

⁴⁴As attorney William Kunstler has pointed out, such a procedure "would not mean that all individuals who follow their own conscience would, or should, go free. A person who is civilly disobedient would have to convince a jury that his moral reasons for disobeying the law were significant -- in fact, that they were more important than the value

of obedience, in this instance. This would certainly be no easy task, but it would provide a means of accommodation, without compromising the law, in cases where the action clearly benefitted society and where no one's interests were harmed." Quoted from Hall, op. cit., pp. 136-137.

⁴⁵I am indebted to David Looman and Margaret Wrensch for bringing this problem to my attention and for discussing it with me. This entire section has also benefitted from discussions I've had with Virginia Tuttle.

⁴⁶A recent example is Federal District Judge Julius Hoffman's conduct in the conspiracy trial of "The Chicago Seven". The verdict against the defendants was overturned by a higher court on the grounds that Hoffman prejudiced the jury against the defendants, among other things.

End