

# *Contents*

## I. LAW AND AUTHORITY

- |                         |    |
|-------------------------|----|
| 1. Legitimate Authority | 3  |
| 2. The Claims of Law    | 28 |

## II. THE NATURE OF LAW AND NATURAL LAW

- |  |     |
|--|-----|
| 3. Legal Positivism and the Sources of Law | 37  |
| 4. Legal Reasons, Sources, and Gaps        | 53  |
| 5. The Identity of Legal Systems           | 78  |
| 6. The Institutional Nature of Law         | 103 |
| 7. Kelsen's Theory of the Basic Norm       | 122 |
| 8. Legal Validity                          | 146 |

## III. INTERNAL LEGAL VALUES

- |                                    |     |
|------------------------------------|-----|
| 9. The Functions of Law            | 163 |
| 10. Law and Value in Adjudication  | 180 |
| 11. The Rule of Law and its Virtue | 210 |

## IV. MORAL ATTITUDES TO THE LAW

- |   |     |
|---|-----|
| 12. The Obligation to Obey the Law                  | 233 |
| 13. Respect for Law                                 | 250 |
| 14. A Right to Dissent? I. Civil Disobedience       | 262 |
| 15. A Right to Dissent? II. Conscientious Objection | 276 |

## V. APPENDICES

- |   |     |
|---|-----|
| 16. The Purity of the Pure Theory   | 293 |
| 17. The Argument from Justice, or How Not to<br>Reply to Legal Positivism | 313 |

- |              |     |
|--------------|-----|
| <i>Index</i> | 337 |
|--------------|-----|



PART I  
LAW AND AUTHORITY



# Legitimate Authority\*

## I. The Paradoxes

There is little surprise that the notion of authority is one of the most controversial concepts found in the armoury of legal and political philosophy. Its central role in any discussion of legitimate forms of social organization and of legitimate forms of political action makes the indefinite continuation of this controversy inevitable. The immediate relevance of the problem of authority to current controversial issues makes a dispassionate study of the subject all the more difficult. But beyond these extrinsic difficulties, the study of the concept of authority has to confront two major problems of intellectual origin: the methodological problem of how to avoid confusing the various quite distinct problems involving the notion of authority and the problem of the paradoxes of authority.

The paradoxes of authority can assume different forms, but all of them concern the alleged incompatibility of authority with reason or autonomy. To be subjected to authority, it is argued, is incompatible with reason, for reason requires that one should always act on the balance of reasons of which one is aware.<sup>1</sup> It is of the nature of authority that it requires submission even when one thinks that what is required is against reason. Therefore, submission to authority is irrational. Similarly the principle of autonomy entails action on one's own judgment on all moral questions. Since authority sometimes requires action against one's own judgment, it requires abandoning one's moral autonomy. Since all practical questions may involve moral considerations, all practical authority denies moral autonomy and is consequently immoral.<sup>2</sup>

\* First published in Richard Bronaugh (ed.), *Philosophical Law* (Westport, Connecticut, 1978). I am indebted to J. E. J. Altham, K. Antley, L. J. Cohen, Philippa Foot, P. M. S. Hacker, and P. H. Nowell-Smith for their critical comments on an earlier version of the essay.

<sup>1</sup> For a version of this principle of reason see Davidson's principle of continence in his 'How Is Weakness of the Will Possible?', in *Moral Concepts*, J. Feinberg (ed.) (Oxford, 1969).

<sup>2</sup> This argument does not apply to theoretical authority. There is nothing immoral

Arguments along these lines do not challenge the coherence of the notion of authority nor do they deny that some people are believed to have authority or actually have *de facto* authority. They challenge the possibility of legitimate, justified, *de jure* authority. Their paradoxical nature derives not from their denial of legitimate authority but from the fact that the denial is alleged to derive from the very nature of morality or from fundamental principles of rationality. Moreover the arguments challenge the legitimacy not only of political authority but of all authority over rational persons.<sup>3</sup> If the very nature of authority is incompatible with the idea of morality and rationality, then those who believe in legitimate authority are not merely wrong or mistaken in one of their moral beliefs. They are committed to an irrational belief or are guilty of a fundamental misapprehension of the concept of morality or of that of authority. This gives these arguments a much greater force. They are, for example, immune from most sceptical arguments. For even if there is no way of distinguishing between right and wrong substantive moral beliefs, at least we can clarify moral concepts and establish relations of entailment and incompatibility between them. If the very concepts of morality and rationality are incompatible with that of authority, then even the sceptic will be able to know that all authority is immoral and submission to it is irrational.

Paradoxically the very force of these arguments is their weakness. Many who may be willing to accept lesser challenges to legitimate authority will be reluctant to accept this most sweeping challenge. Many who may be ready to accept that many authorities are not legitimate, even that no political authority is ever legitimate, will be deterred by the thought that no authority can ever be legitimate. Many who may be ready to concede that those who believe in the possibility of legitimate authority are wrong will shy away from the thought that they are irrational or have no idea what morality is about.

---

in having authorities on how to cook, program a computer, reduce money supply, and so on, so long as one regards them as theoretical authorities only. Submission to theoretical authority may, however, be irrational, for arguments about the conflict between authority and reason are not confined to practical authority.

<sup>3</sup> They may be compatible with authority over small children and some mentally ill people.

It is not my aim to examine the ways in which authority can be defended or attacked. But since the arguments on which the paradoxes are based are said to derive all their force from the analysis of the concepts of authority, morality, and rationality, their examination is relevant to any attempt to clarify the notion of authority. I am concerned here with the nature of authority. I shall try to show why the concept of authority gives rise to the apparent paradoxes and why they are merely apparent. I am not the first to try to dissolve the paradoxes, and it is not part of my brief that all previous attempts to do so have failed. I think, however, that the analysis that follows, even if proving a known truth, does so in a novel way, which shows both the temptation and the fallacies of the paradoxes to best advantage.

## II. A Methodological Detour

Some of the classical authors sought to explain the nature of authority by explaining the way in which people come to accept the authority of individuals or groups. Discussions of the concept were mixed with descriptions of the evolution of society, of conquests, or of social contracts. Modern authors have avoided this confusion, but discussion of the subject is still bedevilled with many methodological confusions. I shall describe briefly four of the common types of explanation and try to point to the lessons to be learned from their shortcomings.

1. The first standard explanation consists in specifying the conditions that are in fact either necessary or sufficient for holding effective (*de facto*) authority. But such explanations fail to elucidate the nature of authority in any way at all. To be sure, it is an important part of social theory to explain under what conditions people can obtain or hold authority, under what circumstances a community is likely to accept the authority of some persons. But they fail altogether to explain what these conditions are for, what it is to have authority or to be in authority.

2. The second type of explanation attempts to elucidate the nature of authority by describing the necessary or sufficient conditions for the holding of legitimate (*de jure*) authority.

This second pattern of explanation seems more promising. According to it the concept of authority is to be explained by explaining how claims to authority can be justified. The force of such explanations is clear. They do not presuppose that claims to authority can in fact ever be justified, but merely point out how they are to be justified. On the reasonable assumption that claims to authority are a way of justifying action, it seems almost inevitable that they differ from other justifications of action by the type of justificative arguments involved. In fact this conclusion is far from inevitable. Justifying claims may differ not only in the nature of the justifying argument invoked but also in the nature of the act justified.

There is a considerable plausibility in the idea that authority is to be explained by reference to the kind of act, for example a claiming, which it justifies. We certainly need authority to perform some actions but not others, and it appears, at least *prima facie*, that to say that one has a certain authority is to indicate that one could be either justified or capable of doing certain actions, without committing oneself in any way as to the nature of that justification.

Here lies the major problem of justificative analyses of authority. None has so far succeeded in delineating the type of argument invocation of which is tantamount to a claim of authority. The fact that there are many different types of authority concerned with virtually every sphere of human activity makes one inclined to give up hope that such a delineation is possible. We can exemplify this difficulty by considering an interesting attempt made recently to provide a justificative analysis of authority. Richard Tuck has suggested that citations of political authority are statements designed to kill criticism of a political action but are not authentic justifications. They are based on the claim that (1) the action proposed or performed is right if somebody performs it; (2) it is neither right nor wrong that the person for whom authority is claimed should be that somebody; (3) that person in fact performed the action or proposes to do so.

Many will share Tuck's belief that nobody has a right to a position of (political) authority and that the only way to justify political authority is by the use of arguments of the type he outlines. But we all know that various writers have thought that

some people are by nature slaves and that those who are by nature free have authority over them. Others have believed in the Divine Right of kings, and there are and were other theories asserting that some people have a right and a duty by nature or by reason to rule. Such people, let us assume, are wrong. But are they also guilty of misusing language, as Tuck's approach suggests?<sup>4</sup> Is the mistake one of moral and political theory, or is it also a mistake about the meaning of words, about the concept of authority?

All other justificative explanations have to overcome the same difficulty. It is not enough to establish that only arguments of a certain type can justify authority. One has to show that claiming authority on any other grounds is a misuse of language.

The criticism of the first two patterns of explanation points to a clear lesson: the analysis of authority cannot consist exclusively of an elucidation of the conditions under which one has either legitimate or effective authority. It must explain what one has when one has authority. This strongly suggests that authority is an ability to perform certain kinds of action.<sup>5</sup> The analysis I have proposed here is meant to vindicate this suggestion.

3. One popular theory that regards authority as ability to perform certain kinds of action identifies effective (*de facto*) authority with power over people. I shall suggest later that to have authority over people is to have normative power. But it is a different notion of power that is involved here. According to it to have power is to have influence, to be able to influence people's actions and their fortunes. A person has effective authority if he is powerful, if he can influence people's fate and their choices or options.<sup>6</sup> Legitimate authority can then be defined as justified effective authority.<sup>7</sup> It is effective authority

<sup>4</sup> Richard Tuck, 'Why Is Authority Such a Problem?' in *Philosophy, Politics and Society*, 4th series, P. Laslett, W. G. Runciman, and Quentin Skinner (eds.) (Oxford, 1972).

<sup>5</sup> My distinction between explanation in terms of types of justification and those in terms of types of actions justified is in itself problematic. Characterizing having authority as ability to perform actions justified by arguments of certain kinds is justificative explanation and not explanation as an ability to perform acts of a certain kind.

<sup>6</sup> For a stimulating discussion of the notion of power as influence see S. Lukes, *Power: A Radical View* (London, 1974).

<sup>7</sup> It is sometimes defined as effective authority accepted by those subject to it or

that should be preserved or obeyed (subject to various conditions and qualifications).

For several reasons, however, it seems that such theories put the cart before the horse. The notion of legitimate authority is in fact the primary one. For one thing not all legitimate authority is effective. Besides (as I will claim shortly), the notion of effective authority cannot be explained except by reference to legitimate authority. Several considerations should be borne in mind.

Though our concern is with practical rather than theoretical authority, an analysis maximizing the similarities between authority for action and authority for belief is, other things being equal, preferable. It seems clear that scientific genius can go unrecognized and that a man who is in fact the greatest authority in a certain field may have very little influence over people's research or their beliefs on issues within his competence.

Parents have authority over their children regardless of whether their children actually acknowledge their authority. Admittedly parental authority is usually recognized by other adults, but that is the wrong sort of recognition from the point of view of a recognition theory, which holds that it is recognition by the subjects that matters. Parental authority does not depend on recognition.

If theoretical authority does not entail recognition or enforcement, then there must be at least some cases of practical authority that also do not entail recognition or enforcement. There are practical authorities whose authority is based entirely on their being theoretical authorities: an expert doctor is an authority not only on the causes of illness but also on their cures. There are experts on the stock exchange and experts on navigation and many others who are authorities for action in their field even though their authority may be unrecognized and unenforced.

I share the belief that a legitimate political authority is of necessity effective at least to a degree. But this is a result of substantive political principles (e.g. that one of the main justifications for having a political authority is its usefulness in securing

---

based on their consent. These facts, however, are relevant, if at all, only to the extent that they show the authority to be a justified one.

social co-ordination, and that knowledge and expertise do not give one a right to govern and play only a subordinate role in the justification of political authority). It is not entailed by a conceptual analysis of the notion of authority, not even by that of the concept of political authority.

The analysis of legitimate authority is not by itself sufficient to explain our notion of authority. A complete account must include an analysis of effective authority as well. Having argued that the notion of legitimate authority does not presuppose that of effective authority,<sup>8</sup> it may be worth pointing out that the reverse is not true. The notion of legitimate authority is presupposed by that of effective authority. A person needs more than power (as influence) to have *de facto* authority. He must either claim that he has legitimate authority or be held by others to have legitimate authority. There is an important difference, for example, between the brute use of force to get one's way and the same done with a claim of right. Only the latter can qualify as an effective or *de facto* authority. But this is a problem that cannot be explored here.

4. Some people hold that authority must be defined by reference to rules: that a person has authority means that there is a system of rules, which confers authority on him. This mode of explanation is in fact a variant of the first and second patterns of explanation and is open to the same basic objection. It substitutes a claim as to when people have authority for a proper explanation of what it is to have authority. It states that people have authority only when it is conferred on them by some rules. But it does not provide any means of deciding which rules confer authority and which do not. Some rules will, it is true, confer authority quite explicitly. They have authoritative, binding formulations (enacted rules), and their authoritative formulations specify that they confer authority on a certain person. But the proposed definition does nothing to illuminate their meaning and effect.

The claim that all authority is conferred by rules is itself debatable. It is difficult to maintain that when a member of the

<sup>8</sup> This should not lead one to confuse authority with the right to have authority. A person may be entitled to have authority and yet not have it: he may be entitled to it and entitled to have it conferred on him by Parliament and yet not have it until it is so conferred on him.

public assumes authority in an emergency (for example, a fire in a theatre) his authority derives from any rules. It is not, however, my purpose here to discuss the ways authority can be acquired or defended. There are other objections to definitions of this type. Unless properly qualified they entail contradictions. If there are two systems of rules according to one of which a certain person has authority whereas according to the other he does not, then he both has and does not have authority. To avoid such a contradiction the proposed definition must be relativized. It cannot be taken to be a definition of having authority but of the relativized notion of having authority according to  $s$  where  $s$  is some system of rules. The relativized notion of authority, however, severs the connection between authority and practical reason.

Authority is a practical concept. This means that questions of who has authority over whom are practical questions; they bear on what one ought to do. In other words statements that some persons have authority may serve as premisses in practical inferences. The explanation of authority must explain the practical import of the concept. It must explain how it is capable of figuring in practical inferences.

What one ought to do depends on who has authority in a non-relativized sense. That a person has authority according to some system of rules is, in itself, of no practical relevance. Just as one can draw no conclusions as to what ought to be done from the fact that according to a certain person authority is vested in Parliament, so one cannot draw any such conclusions from the mere fact that according to some rules authority is vested in Parliament. Certain further assumptions may entail that if according to someone Parliament has authority, then Parliament does have authority. Similarly, further assumptions may allow a move from a statement of authority according to some rules to a non-relativized statement of authority. It would be a mistake, however, to build those further conditions into one's definition of 'authority according to some rules' so as to make the move to a non-relativized statement of authority always possible. The whole purpose of talk of relativized authority is to block the possibility of such a move, unless further assumptions are available. We need such a device to be able to talk of the views of other people about authority,

of the situation according to rules accepted in some societies or proposed by some people. To do this we have in talk of relativized authority a way to refer to what those people or societies accept or propose as legitimate authority without endorsing those views. We simply state what authority is had by whom from a certain point of view.<sup>9</sup> In some circumstances the fact that some people hold certain views or endorse certain rules is enough to invest a person with authority. In others it is not. The move from a relativized statement of authority to a non-relativized one is never automatic and is not always possible.

These considerations suggest that the non-relativized notion is the primary one. The relativized notion is useful because it reveals the views of people or societies concerning non-relativized authority. Its explanation presupposes the non-relativized notion, which does not presuppose it. Our task then is to explain the notion of legitimate non-relativized authority in a way that shows its relevance to practical reasoning.

### III. The Simple Explanation

Several authors have analysed authority along the lines I have just suggested. On the whole there is a large measure of agreement between them but they differ greatly in important details. Robert Paul Wolff, to take one well-known example, says that 'authority is the right to command, and correlatively, the right to be obeyed'.<sup>10</sup> His definition is essentially sound, but it is both inaccurate and not perspicuous. It is inaccurate, for authority is a right to do other things as well. It can be a right to legislate, to grant permissions, to give authoritative advice, to adjudicate, and so forth. It is wrong to regard all these as commanding. Wolff's definition is not perspicuous since the notion of a right is even more complex and problematic than that of authority.

To be useful, the analysis must be made in terms of relatively simple concepts. From this point of view I think that the best existing explanation of authority is that offered by John Lucas: 'A man, or body of men, *has authority* if it follows from his saying "Let X happen", that X ought to happen.'<sup>11</sup> This definition

<sup>9</sup> The use of expressions such as 'according to law' is just one way of indicating that the statement is made from a point of view only. Often the context of utterance suffices to indicate this.

<sup>10</sup> Robert Paul Wolff, *In Defense of Anarchism* (New York, 1970), p. 4.

<sup>11</sup> John Lucas, *The Principles of Politics* (Oxford, 1966), p. 16.

is both perspicuous and general. It applies to all types of practical authority over persons and not merely to political authority. It makes clear that one can exercise authority not only by commanding but in other ways as well.<sup>12</sup>

Lucas explains authority as an ability to perform an action, and he regards the relevant action as that of changing the normative situation. I shall assume that if  $X$  ought to  $\phi$  then he has a reason to  $\phi$  and that if he has a reason to  $\phi$  then he ought to  $\phi$ . On this assumption Lucas's definition entails that a person has authority if his saying, 'Let  $X$  happen', is a reason for  $X$  to happen. This sounds somewhat incongruous. The reason is that Lucas's definition does not make clear that the authority he is defining is authority over persons. To make this assumption clear we can further amend his definition and say that  $X$  has authority over  $Y$  if his saying, 'Let  $Y \phi$ ', is a reason for  $Y$  to  $\phi$ . Let us call this the simple analysis.

Two comments are in place here. First, I do not claim that authority can be explained only in terms of reasons. The preference for a reason-based explanation is motivated by the belief that reasons provide the ultimate basis for the explanation of all practical concepts, namely, that all must be explained by showing their relevance to practical inferences. The preference for reason-based explanations of authority is one for trying to show the role of authority statements in practical reasoning directly rather than through the mediation of other concepts (such as rights). Second, a great variety of things are called reasons. That it rains, for example, is a reason for carrying an umbrella. So is the fact that one wants to be outside and not get wet. But in a perfectly straightforward sense both are just parts of one reason. We can distinguish between partial reasons and the complete reasons of which they are parts. It is in terms of complete reasons that the attempt to analyse authority will be made.<sup>13</sup> I shall argue below that the simple explanation fails

<sup>12</sup> His definition confines the exercise of authority to the use of optatives. It seems that authority can also be exercised through non-verbal behaviour and communication. But I am not concerned here with the examination of the different ways in which authority can be exercised.

<sup>13</sup> A complete reason on this account is not necessarily one whose justification is self-evident. That I promised to perform a certain action is, I believe, a complete reason to perform it. But, of course, one may well be challenged to justify such beliefs and must be ready to justify them. One must defend, in other words, the belief that promises are reasons for actions. For further

to distinguish adequately between intentional and non-intentional exercise of authority and that it does not pay attention to the distinction between being an authority and having authority. It also overlooks the fact that one needs authority to grant permissions and to confer powers. But first I shall examine some more far-reaching objections to it.

#### IV. First Objection to the Simple Explanation

Is the utterance of an authority an absolute or a *prima facie* reason for doing as it demands? If we assume that it is an absolute reason, then it seems very unlikely that there are any legitimate authorities, and there seem to be very few *de facto* authorities. I, for example, believe that I am justified in taking the advice, commands, or rules issued by some people as reasons for action, but I cannot see that it is ever right to take anybody's word as an absolute reason to be followed under all circumstances. It seems to me that this is a widely shared view and that most people hold that under certain circumstances the instructions of authority need not be followed. Hence if authority is explained in terms of ability to issue absolutely binding instructions, then there seem to be very few recognized authorities in the world and none which is legitimate.

But are authoritative utterances *prima facie* reasons? Compare an order with a request and both with advice. All three are identified by the attitudes, beliefs, and intentions of their source, not by the way they are received by their addressee. The fact that one was ordered or requested or advised to take a certain action may be a reason to take it, and it may be held by the addressee to be so. Under different circumstances the fact that such utterances were made is no reason for action or is not held to be one. From the addressee's point of view there is no necessary difference between being ordered, being requested, and being advised except that they entail or imply different intentions, beliefs, or attitudes in the person issuing them.

One such difference is that the primary intention in advising is to convey information about what is morally right or wrong, what is lawful or unlawful, in one's interests or not, and so forth

---

discussion of the presupposition underlying my use' of 'reasons' see *Practical Reason and Norms* (2nd ed, Oxford, 1999), ch. 1.

or just about brute facts. If there is an intention to influence the addressee (and there need not be one) then it is to influence him by making him aware of the situation (for example, that he ought to  $\phi$  or that  $\phi$ -ing will secure the greatest income possible in the circumstances).

In short the adviser must intend his giving the advice to be taken as a reason to believe that what he says is true, correct, or justified. But he does not necessarily intend it to be taken as a reason for action, even though it may be the case that his giving the advice is a valid reason for action for the recipient.<sup>14</sup>

Requesting and ordering, on the other hand, entail intending that the act of requesting or ordering be taken as a reason to perform the act ordered or requested. What then is the difference between them? One such difference is relevant to our purpose. Suppose that a man makes a request and is told in reply that his request was considered, but on balance it was found that the reasons against the action requested overrode those for it including the request itself. He will no doubt be disappointed, but he will not feel that his request was disregarded. He has nothing to complain about. He must concede that whatever his hopes, he intended no more than that the action be taken on the balance of reasons, his request being one of them.<sup>15</sup> This is not so if he gave an order. A man who orders someone else does not regard his order as merely another reason to be added to the balance by which the addressee will determine what to do. He intends the addressee to take his order as a reason on which to act regardless of whatever other

<sup>14</sup> These remarks apply most naturally to 'advise that  $P$ '. Are they true of 'advise to  $\phi$ '? This is a moot point, but it seems to me that 'I advise you to apply to Balliol', when used to advise, is used to make the same statement as is often made by 'Balliol is your best choice', or by 'On balance I think applying to Balliol is preferable to the alternatives'. 'Advise to  $\phi$ ', is reducible to 'advise that  $P$ '. But I shall not argue for this here.

<sup>15</sup> Some requests, such as pleas and beggings, intend also to induce new spontaneous desires to accede to them. A 'pure' request appeals to existing sympathies and to reason as sufficient to make it a reason for action for the addressee. It is wrong, however, to think of requests as merely intended to 'activate' a reason the addressee has (his sympathy or recognition of duty) by informing him of a need or a desire for the object of the request. Information of somebody's need, even if conveyed with the intention to move the addressee to action by activating existing reasons, does not amount to a request. One requests only when the act of communication itself is meant to constitute a reason. In special cases one has a right that one's requests be granted. There the divide between requests and commands is thin indeed and does on occasion boil down to the choice of language.

conflicting reasons exist (short usually of an emergency or other extreme circumstances).<sup>16</sup>

It may seem that the explanation of this difference is that a man issuing an order always intends it to be a very weighty reason and that is not always the case when people make requests. Apart from the fact that some requests are made with such an intention, this explanation seems unsatisfactory because it relies on an alleged difference in degree. If the difference indicated is the one crucial to the distinction between orders and requests, then it is unlikely to be one of degree only.

The crucial point of this objection is that one requires authority to be entitled to command but one does not need authority to be entitled to request. My point is not that everyone is entitled to request. Whether this is the case is a moral, not a conceptual, question. There is nothing in the concept of a request that entails that everyone is entitled to request. My point is that the fact that one is entitled to request does not entail that one has authority over the addressee of the request. Contrariwise, that one is entitled to command entails that one has authority over the addressee of the command. A request made by a person entitled to make it is a valid (*prima facie*) reason for its addressee. Similarly a command issued by a person entitled to issue it is valid. If we are to say no more than that a valid command is a reason for its addressee, then we fail to explain the difference between a command and a request and the reason for which only entitlement to the first entails having authority. To say that a valid command is a weightier reason than a valid request is both false and inadequate as an explanation of a distinction that is not merely one of degree.

## **V. Second Objection to the Simple Explanation**

The first objection was based on an argument to the effect that if authority is ability to change reasons by certain utterances,

<sup>16</sup> This statement has to be modified if it is to apply to orders (and requests) made in an institutional setting (such as the army). There conventional ways of commanding (or requesting) normally used to facilitate the identification of the agent's intentions are hardened and the agent's action invoking the convention counts as a command regardless of his true intentions. Such conventions presuppose that typically they will be used with the appropriate intention. When this presupposition fails they can no longer be regarded as conventions of commanding (or requesting).

then the utterances of authority are more than *prima facie* yet less than absolute reasons, which is an impossibility. The second objection is based on an argument to the effect that the utterances of (legitimate) authority though often reasons for action need not always be so. It consists of an appeal to our intuition based on a counter-example. Consider the following situation. I am driving my car in flat country with perfect visibility and there is no other human being, animal, or car for miles around me. I come to a traffic light showing red. Do I have any reason to stop? There is no danger to anyone and whatever I do will not be known to anyone and will not affect my own attitude, feelings, or beliefs about authority in the future. Many will say that there is not even the slightest reason to stop at the red light in such circumstances.<sup>17</sup> They insist that this in no way contradicts their acknowledgement of the legitimate authority of those who made the traffic regulations. This example seems sufficient to convince one that in this case or a similar case the utterances of authority can be held to be legitimate without holding them to constitute reasons for action.

I think that these objections are sufficient to undermine the simple explanation and yet the simple explanation is right in its basic insight—that authority is ability to change reasons for action. Both the simple explanation and the objections to it are based on an over-restricted view of reasons for action. I shall argue that if we regard authority as ability to change a certain type of reason, then the objections can be easily answered.

## VI. Normative Power

Consider any situation in which an authority instructs a subordinate to follow the instruction of another whose authority does not derive from that of the first. A father telling his son to obey his mother is such a case. It differs from the father telling the son to obey his nanny, since the nanny's authority derives from that of the parents. An instruction to obey the nanny is, we may assume, her only source of authority. The mother's instructions are in any case authoritative. They are reasons for

<sup>17</sup> Further assumptions are required to show that I know that I have no reason to stop. For my purpose it is enough to establish that in fact I have no reason to stop. But in many situations these further assumptions obtain, and it is also true that for all I know there is no reason to stop.

action for the son. So are the father's instructions. His instruction to obey the mother is, therefore, a reason to act for a reason. It is a reason to act on the mother's instruction, which is itself a reason anyway. I shall call a reason to act for a reason a positive second-order reason. There are also negative second-order reasons, that is, reasons to refrain from acting for a reason. I shall call negative second-order reasons exclusionary reasons. To get an example of an exclusionary reason we need only reverse the father's instruction and assume that he orders his son not to act on his mother's orders.<sup>18</sup> Now the son has a reason for not acting on a reason.

There is one important point to bear in mind concerning second-order reasons: they are reasons for action, the actions concerned being acting for a reason and not acting for a reason. If  $P$  is a reason to  $\phi$  then acting for the reason that  $P$  is  $\phi$ -ing for the reason that  $P$ . Not acting for  $P$  is not  $\phi$ -ing for the reason that  $P$ . This is compatible with  $\phi$ -ing for some other reason as well as with not  $\phi$ -ing at all. I am not assuming that whenever one fails to act on a reason one does so intentionally. One may fail to act on a reason because one does not know of its existence. These clarifications make it plain that in the examples I am assuming that in telling his son to obey his mother, the father tells him not merely to do what she tells him to do but also to do it for the reason that she tells him so. Similarly I am assuming that when the father tells his son not to obey his mother, he is not telling him never to do what his mother tells him to do but merely never to take her instructions as reasons for action.

Sometimes a person may have a reason for performing an action and also a reason for not acting for certain reasons against that very action. The son, in our example, may know that his only coat is ugly. This is a reason against wearing it. It conflicts with his mother's instruction to wear a coat when he goes out at night. But the reason against wearing the coat is reinforced indirectly by the father's order to disregard the mother's instruction. In this and many other cases the fact that is a reason (the father's order) for disregarding certain reasons (the mother's instruction) for  $\phi$ -ing (wearing the coat) is

<sup>18</sup> I am assuming that he is entitled to make such an order—a matter that can be subject to dispute but that need not detain us here.

different from any fact that is a reason (the coat's ugliness) for not  $\phi$ -ing. But sometimes the same fact is both a reason for an action and an (exclusionary) reason for disregarding reasons against it. I shall call such facts protected reasons for an action.

I will define normative power as ability to change protected reasons. More precisely, a man has normative power if he can by an action of his exercise normative power. An act is the exercise of a normative power if there is sufficient reason for regarding it either as a protected reason or as cancelling protected reasons and if the reason for so regarding it is that it is desirable to enable people to change protected reasons by such acts, if they wish to do so.<sup>19</sup>

I shall assume that power is used by making what I shall call 'power-utterances'. There are three ways in which the power-holders can change protected reasons that are important to our purpose. The first is by issuing an exclusionary instruction, that is, by using power to tell a person to  $\phi$ , the power-utterance is a reason for that person to  $\phi$  and also a second-order reason for not acting on (all or some) reasons for not  $\phi$ -ing. Exclusionary instructions are, therefore, protected reasons. The second way of exercising power is by making a power-utterance granting permission to perform an action hitherto prohibited by an exclusionary instruction. I shall call such permissions 'cancelling permissions' for they cancel exclusionary reasons. The third form of using power is by conferring power on a person. This does not in itself change protected reasons, but it enables a person to change them. The power a person has can be restricted in many ways—in the way it can be exercised, the persons over whom it is held, the actions with respect to which the power-holder can make power-utterances, and so forth.

Given these clarifications it is evident that there is a close relation between normative power and authority. On the simple explanation of authority power is a special case of

<sup>19</sup> Cf. my 'Voluntary Obligations and Normative Powers', in *Proceedings of the Aristotelian Society*, Supp. Vol. 46 (1972), 79–102, and my *Practical Reason and Norms*, section 8, where the notion is more fully explained as well as given a more general definition. In this book, however, I erred in suggesting that a normative power is ability to change exclusionary reasons. Comments made by Philippa Foot forced me to realize that it is ability to change protected reasons (or exclusionary permissions). All mandatory rules are protected reasons.

authority. Authority is ability to change reasons. Power is ability to change a special type of reasons, namely protected ones. However, in the light of the objections to the simple explanation, we should regard authority basically as a species of power. To provide a comprehensive defence of this view requires showing that rules and commands are protected reasons and that all authoritative utterances are power-utterances. This is not a task that can be accomplished in this essay. Instead I shall try to provide a persuasive case showing, first, that the two objections fail against this view and, second, that the paradoxes can be explained away by it.

### VII. Power and Authority

We should distinguish between authority over persons and authority to perform certain actions. The two overlap but are distinct notions. Everyone who is an authority has authority over people, but not everyone who has authority is an authority. The difference is not of great philosophical moment, but its neglect can be a source of endless confusion. A person is an authority if he has relatively permanent and pervasive authority over persons, that is, either authority over a large group of people or with respect to various spheres of activity, or both.

Since power is the ability to change protected reasons for action and as reasons for action are reasons for some persons or others, we can divide powers into powers over oneself and powers over others. The most important species of power over oneself is the power to undertake voluntary obligations. Power over others is authority over them. There is one exception to this characterization. Sometimes we say that a person has authority over himself. This is a degenerate case of authority: an extension by analogy from the central cases of authority over others. It is interesting to note that when speaking of a person's authority over himself, we always refer to his power to grant himself permissions or powers. We never refer thus to one's power to undertake voluntary obligations.

One of the main obstacles to an analysis of authority is the frequent failure to distinguish between authority to perform an action and authority over persons.<sup>20</sup> A person has authority to

<sup>20</sup> This confusion undermines much of Richard Tuck's analysis in 'Why Is Authority Such a Problem?' (n. 4 above).

perform an action if he has been given permission to perform it or has been given power to perform it by somebody who has power to do so. Thus, I have authority to open your mail if the censor has given me permission to do so, assuming that he has power to do so.

My authority to open your mail is not authority over you. I cannot change your normative situation in any way though the censor changed it by giving me the authority to open your mail, thereby diminishing your right to privacy. I can also have authority to sign cheques in your name, which is a power I have because you gave it to me. This last example shows that the source of a person's authority to perform an act must have power to confer it, but he need not have authority over the person on whom he confers authority. To give me authority to sign cheques in your name you need power, but that you have such a power does not entail that you have any authority over me. It may well be that you have none.

The authority to act is, however, closely related to authority over persons, albeit in a somewhat more indirect way. When we consider the cases in which we are granted permissions or powers, it is evident that not all of them can be described as having authority to act. Only when the interests of another person will be affected by the act do we speak of it as authorized. But this is not a sufficient condition. I am permitted to open a supermarket that will lead to someone's bankruptcy, and yet it is not the case that I have authority to open the supermarket. I am permitted to do so simply because there is not and never was a prohibition against doing so. One has authority to do only those things that one is given permission to do by somebody who has authority over the person whose interests are affected.

We can now define *X has authority to  $\phi$*  as: there is some *Y* and there is some *Z* such that,

- (1) *Y* permitted *X* to  $\phi$  or gave him power to do so
- (2) *Y* has power to do so
- (3) *X*'s  $\phi$ -ing will affect the interests of *Z* and *Y* has authority over *Z*.<sup>21</sup>

<sup>21</sup> Normally *X*, *Y*, and *Z* will be different persons. A person may, however, give authority to another to affect his own interests or give himself authority to affect the interests of another.

### VIII. Refuting the Objections

The distinctions between being an authority and having authority and between authority to act and authority over persons are not directly connected with my claim that authority over persons, the basic concept of the three, is a species of normative power. These distinctions must be preserved by any account of authority. The advantage of the power analysis of authority is that it successfully meets the objection to the simple explanation and dissolves the paradoxes of authority.

Take the first objection first. What is the difference between orders and requests? Authority over persons is ability to change protected reasons for their actions. In most discussions of the concept of authority attention is focused on issuing orders and laying down rules as the standard manifestations of authority. These are, indeed, the standard cases of the intentional invocation of authority. It is, however, important to see that authority can be exercised without the person having authority intending to invoke it. This is true of political as well as of other kinds of authority and is a very important channel for the influence of authority to be felt. I shall take advice given by authority as a typical case in point.

Advice, whatever the hopes of the adviser may be, is given with the intention that its utterance will be taken as a reason for belief, not for action.<sup>22</sup> But the recipient of the advice may regard it as both a reason for action and an exclusionary reason for disregarding conflicting reasons. Consider the standard grounds for seeking advice. These are usually to gain information relevant to the solution of some practical problem facing one or for comparing one's own evaluation of the weight and importance of various factors with that of other people as a means of checking one's own views and calculations. But sometimes advice is sought for entirely different reasons. A person may be faced with a problem involving considerations concerning which he has little knowledge or understanding. He may turn to an expert, to an authority, for advice despite the fact that he has no way of assessing the reasons pointed out by the

<sup>22</sup> Advice differs from other cases of conveying information primarily in being given with the belief that it is or may be relevant to an actual or hypothetical question facing the recipient in his own estimate or in the opinion of the adviser and in being either intended by the adviser or expected by the recipient to be taken into account in the resolution of that practical problem.

authority against other conflicting reasons of which he may be aware. He may decide to follow the advice given without trying to work out whether it indicates reasons that tip the balance. If he does so he is in fact excluding all the conflicting reasons of which he is aware from his considerations. He is regarding the advice both as a reason to perform the action he was advised to perform and for not acting on conflicting reasons. A person may be justified in holding the advice he received to be a protected reason, even though advice is not given with an intention to be taken as a protected reason.

Orders, on the other hand, are given with the intention that their addressees shall take them as protected reasons. Many people can give an order without being entitled to do so. They are entitled to do so only if they have authority (power) over the addressee with respect to the subject-matter of the order. The order may be a valid first-order reason for performing the act even if it is not a valid exclusionary reason not to act on conflicting reasons, and it may be both even though the person who issued it has no authority to do so. But it always is a valid first-order reason and an exclusionary reason if he has the authority to give it.

Exclusionary reasons may exclude action for all or only for some kinds of the conflicting reasons. Exclusionary reasons differ in scope, that is, in the extent to which they exclude different kinds of conflicting reasons. Therefore, to maintain that orders are both first-order and exclusionary reasons is not tantamount to maintaining that they are absolute reasons. They may not exclude certain conflicting reasons, and when this is the case one must decide what to do on the balance of the non-excluded first-order reasons, including the order itself as one *prima facie* reason for the performance of the ordered action.

What then is the difference between an exclusionary reason and a first-order reason of a weight sufficient to override all the conflicting reasons that are excluded by the exclusionary reason and no others? There are two answers to this crucial question. First, exclusionary reasons exclude by kind and not by weight. They may exclude all the reasons of a certain kind (such as considerations of economic welfare), including very weighty reasons, while not excluding even trivial considerations belonging to another kind (such as considerations of

honour). Second, regardless of the different impact of exclusionary and weighty reasons on what ought to be done, all things considered, they also differ in the way we view them. Some facts are weighty reasons overriding conflicting reasons; others are not to be compared with conflicting reasons. Their impact is not to change the balance of reasons but to exclude action on the balance of reasons.

This difference in function, regardless of any possible difference in what ought to be done, all things considered, explains the difference between orders and requests. Valid orders are not necessarily more weighty or important reasons than valid requests. There could be orders that exclude few conflicting considerations and that do not exclude and may be overridden by certain requests. A request may be a reason of sufficient weight to justify sacrificing one's life. The difference is not in importance but in mode of operation. A request is made with the intention that it shall be taken as a reason for action and be acceded to only if it tips the balance. Orders are made with the intention that they should prevail in certain circumstances even if they do not tip the balance. They are intended to be taken as reasons for excluding certain others that may tip the balance against performing that action.

Put it another way. For every order, if we know what the person who issued it thinks is the correct outcome of all possible practical conflicts in which it may be involved, we can ascribe to him the view that his order has just the weight that would justify all those consequences. But in doing this we have not advanced at all towards explaining the difference between orders and requests in general. This can be done if there is a certain constant weight, or range of weights, that is characteristic of all orders and distinguishes them from requests. But in fact both orders and requests span the whole range of possible weights both in the eyes of those who issue them and of others. Having assumed that the difference between them lies in their practical implications, I submit that it consists in the fact that orders but not requests are protected reasons.

There is a minimum that an order must exclude to be an order. It must at least exclude considerations of the recipient's present desires. Often orders exclude much more besides, but never do they exclude less. In appropriate circumstances one

may be able to justify one's not having followed an order on the ground that the order was not intended to apply to the case. It was never intended, one could claim, that one should obey even if it turns out that there was a strong moral reason for not doing so or if obeying would severely damage the recipient's interests or be unlawful.<sup>23</sup> When such considerations amount to a justification and lead the agent not to follow the order, he cannot be said to have obeyed it but neither did he disobey it. It was not intended that he should follow it in such circumstances. But it is never a justification that the agent had a desire, however strong, for something inconsistent with his following the order. Many parents' orders come near to the minimum exclusion, being intended to exclude only consideration of the child's present desires in order to avoid argument on what is best, given his strong desire to perform or to avoid some action. But parents' orders often exclude a consideration of the child's own interests and may exclude much more besides.

That is the explanation of the fact that it is more presumptuous to order than to request. If you request you submit yourself to the addressee's judgment on the balance of reasons, while at the same time trying to add a reason on one side of that balance. But one who commands is not merely trying to change the balance by adding a reason for the action. He is also trying to create a situation in which the addressee will do wrong to act on the balance of reasons. He is replacing his authority for the addressee's judgment on the balance.

Similar considerations refute the second objection which denies that authoritative instructions are always reasons for action. There is a sense in which if one accepts the legitimacy of an authority one is committed to following it blindly. One can be very watchful that it shall not overstep its authority and be sensitive to the presence of non-excluded considerations. But barring these possibilities, one is to follow the authority regardless of one's view of the merits of the case (that is, blindly). One may form a view on the merits but so long as one follows the

<sup>23</sup> This is not always an acceptable justification. Many legal orders, for example, are given with the intention that they shall be followed in the face of conclusive moral reasons against performing the ordered act. It may be that one should never regard such orders as binding and that no one has legitimate authority to issue such orders. But this in no way affects my point that such orders are often given. In the intention of the man giving the order it is to be taken as excluding even moral considerations.

authority this is an academic exercise of no practical importance. We can go further than that and say that sometimes the very reasons that justify the setting up of an authority also justify following it blindly in a stronger sense—that is, following it without even attempting to form a judgment on the merits. This is the case, for example, with some traffic regulations. We all know the benefit from allowing traffic lights to regulate one's action rather than acting on one's own judgment. But we tend to forget that a significant part of the benefit is that we give up attempting to form a judgment of our own. When I arrive at a red traffic light I stop without trying to calculate whether there is, in the circumstances, any reason to stop. From our vantage point we have invented an example in which the question does not arise since the answer (there is no reason) is plain. But for the man in our example the question does arise; he has to discover whether there is no reason to stop. And if he is to inquire in this case, he has to inquire in many other cases. For us it looks ridiculous to hear him say, 'I am bound to follow authority regardless of the merits of the individual case', for we know in advance what the merits are and forget that he has to find that out, and not only now but in many other cases as well. Only when it is justified to prevent this, is it also justified to accept authority in this respect, even if once in a while this makes one look ridiculous to the gods.

### **IX. Dissolving the Paradoxes**

The last remarks in response to the objection to the simple view of authority help to explain both the force of the paradoxes and the way to overcome them. The paradoxes, I should hasten to say, pose no problem to the simple view of authority. On this view, the commands of legitimate authority are facts of the world that are reasons for action. They are essentially like the weather and the stock exchange in being facts that are reasons for certain actions and against others. One no more abandons reason or forfeits one's autonomy if one follows the commands of authority than if one follows trends on the stock exchange.

This solution not only gets rid of the paradoxes but also presents them as simple mistakes. If, however, authority over persons is normative power over them, then we can explain the temptation that the paradoxes present without succumbing to

it. I shall examine the paradoxes as presented by Robert Paul Wolff: 'Men', he says 'can forfeit their autonomy at will. That is to say, a man can decide to obey the commands of another without making any attempt to determine for himself whether what is commanded is good or wise'.<sup>24</sup> Whatever the relevance of the second sentence to the problem of autonomy, it is true that accepting authority inevitably involves giving up one's right to act on one's judgment on the balance of reasons. It involves accepting an exclusionary reason.<sup>25</sup>

Wolff is anxious to emphasize that his view does not require people to disregard orders and commands altogether. The following shows both the strength and the weakness of his position:

For the autonomous man, there is no such thing, strictly speaking, as a command. If someone in my environment is issuing what are intended as commands, and if he or others expect those commands to be obeyed, that fact will be taken account of in my deliberations. I may decide that I ought to do what the person is commanding me to do, and it may even be that his issuing the command is the factor in the situation which makes it desirable for me to do so. For example, if I am on a sinking ship and the captain is giving orders for manning the lifeboats, and if everyone else is obeying the captain because he is the captain, I may decide that under the circumstances I had better do what he says, since the confusion caused by disobeying him would be generally harmful. But insofar as I make such a decision, I am not obeying his command; that is, I am not acknowledging him as having authority over me. I would make the same decision, for exactly the same reasons, if one of the passengers had started to issue 'orders' and had, in the confusion, come to be obeyed.<sup>26</sup>

Wolff is making two valid and important points here.

(1) Because an order is always given with the intention that it be taken as both an exclusionary reason and a first-order reason, its addressee has more options than either to disregard the order altogether or to obey it as he was intended. He may hold it to be a valid first-order reason, given the circumstances of its utterance, while denying that it is an exclusionary reason.

(2) This means that an anarchist can reject the legitimacy of all authority while giving some weight to the instructions of *de*

<sup>24</sup> Wolff, *In Defense of Anarchism* (n. 10 above) p. 14.

<sup>25</sup> Wolff is wrong in saying that accepting authority involves giving up the right or the attempt to form a judgment on the balance of reasons. Only action on that judgment is excluded (if it involves relying on excluded reasons which are not overridden).

<sup>26</sup> *Ibid.* pp. 15–16.

*facto* authorities. He can take such instructions to be first-order reasons without conceding the legitimacy of the authority. For it is only by acknowledging that such instructions are also valid exclusionary reasons that one accepts the legitimacy of the issuing authority. Only such an acknowledgement amounts to submission to authority for only it contains the necessary element of the denial of one's right to act on one's own judgment on the merits.

Reformulating Wolff's contentions in this way does show that there is more in them than simple confusions. He sees correctly that legitimate authority involves a denial of one's right to act on the merits of the case. But the reformulation also shows where he went wrong. He tacitly and correctly assumes that reason never justifies abandoning one's autonomy, that is, one's right and duty to act on one's judgment of what ought to be done, *all things considered*. I shall call this the principle of autonomy.<sup>27</sup> He also tacitly and wrongly assumes that this is identical with the false principle that there are no valid exclusionary reasons, that is, that one is never justified in not doing what ought to be done on the balance of first-order reasons. I shall call this the denial of authority.

This confusion is natural if one conceives of all reasons as essentially first-order reasons and overlooks the possibility of the existence of second-order reasons. If all valid reasons are first-order reasons then it is a necessary truth that the principle of autonomy entails the denial of authority, for then what ought to be done all things considered is identical with what ought to be done on the balance of first-order reasons. But since there could in principle be valid second-order reasons, there is nothing in the principle of autonomy that requires the rejection of all authority.

The question of the legitimacy of authority takes the form that it was always assumed to take: an examination of the grounds that justify in certain circumstances regarding some utterances of certain persons as exclusionary reasons. There is no short cut that will make such an inquiry redundant by showing that the very concept of legitimate authority is incompatible with our notion of rationality or morality.

<sup>27</sup> It is clear that this principle of autonomy is not really a moral principle but a principle of rationality.

## The Claims of Law

Can the very abstract analysis of authority in the previous essay be applied to the law? Do we have any reason to expect the notion of authority to play an important role in our understanding of the law or in shaping our attitude to it? Two preliminary objections must be tackled first.

(1) It is a popular view that the law enjoys *de facto* or effective authority. Its analysis involves these concepts but not necessarily that of a legitimate authority. This is a mistake. To hold that a government is *de facto* government is to concede that its claim to be government *de jure* is acknowledged by a sufficient number of sufficiently powerful people to assure it of control over a certain area. A person has effective or *de facto* authority only if the people over whom he has that authority regard him as a legitimate authority. This would usually, though not necessarily always, imply that he claims legitimate authority for himself. It is enough that others regard him as a legitimate authority. This condition may have to be weakened. It may be unnecessary for the relevant population genuinely to believe that the person having effective authority is a legitimate authority; it is enough if, for whatever reasons, they avow such a belief. Furthermore, we could well find situations in which only some of the population acknowledge the authority of the relevant person, provided they are sufficiently prominent or powerful to enable him to impose his rule on the others. The reasons for which the relevant population accepts the authority of a person vary and do not belong to the analysis of the concept of a *de facto* authority. The population may acknowledge the legitimate authority of a person on the ground that he is the person in control (and that any attempt to change things will lead to chaos, etc.). This should not lead to the conclusions that control is enough and that acceptance of a claim to legitimate authority is not a logically necessary condition for the existence of a *de facto* authority. We need not concern ourselves with the precise conditions for being an effective authority. They probably differ with respect

to different types of effective authorities: legal, political leadership, ecclesiastical, etc. A common factor in all kinds of effective authority is that they involve a belief by some that the person concerned has legitimate authority. Therefore, the explanation of effective authority presupposes that of legitimate authority.<sup>1</sup>

(2) Authority was analysed in the context of a person in authority and his authoritative utterances. Such an analysis could in principle apply to a legislator and his acts of enactment. But not all law is enacted. Customary rules can be legally binding. Can they be authoritative despite the fact that they are not issued by authority? It is possible to talk directly of the authority of the law itself. A person's authority was explained by reference to his utterances: he has authority if his utterances are protected reasons for action, i.e. reasons for taking the action they indicate and for disregarding (certain) conflicting considerations. The law has authority if the existence of a law requiring a certain action is a protected reason for performing that action; i.e. a law is authoritative if its existence is a reason for conforming action and for excluding conflicting considerations. 'Reason' here means a valid or justifiable reason, for it is the legitimate authority of the law which is thus defined. The law enjoys effective authority, as we saw it, if its subjects or some of them regard its existence as a protected reason for conformity.

These clarifications show that it makes sense to refer to the authority of the law and that if its analysis involves the notion of effective authority it also involves that of legitimate authority. This may seem sufficient to show that the notion of authority is inextricably tied up with that of law. It is indeed plain that to determine our proper attitude to the law we must examine whether the law has authority over us which we should acknowledge. But need we use the concept of authority in explaining the nature of law? Does the concept play a part not only in determining our obligation to the law but also in our understanding

<sup>1</sup> This view is in fact incorporated in most current analyses of authority. If it appears at first sight controversial, this is only because it is implicit rather than explicit in many popular accounts of *de facto* authority. It is generally agreed that authority cannot be based on force or threat of force alone; that it is dependent on influence, acceptance, etc. Further analysis will show that these notions presuppose a claim to authority and its recognition.

of it? The answer is affirmative, for it is an essential feature of law that it claims legitimate authority.

This may sound unduly mysterious, but it can in fact be given fairly precise meaning. We should concentrate on the circumstances in which non-conforming behaviour is or is not a breach of law. We need to establish first that the law claims that the existence of legal rules is a reason for conforming behaviour. This should not be confused with the false claim that the law requires conformity motivated by recognition of the binding force, the validity of the law. It is a truism that the law accepts conformity for other reasons (convenience, prudence, etc.). If pronouncements from the Bench may be taken as its mouthpiece,<sup>2</sup> it sometimes welcomes reliance on other (for example, independent moral) reasons for doing that which the law requires. Suppose, however, that—apart from the law itself—there are insufficient reasons for doing as the law requires. Suppose that on balance (excluding the existence of the law from the balance) one ought not to perform the required act, and that that is the view of both the law's subjects and of the judge. Failure to perform the act in such circumstances is a breach of the law. Does this mean that the law requires action against reason? No, it merely means that the law holds itself, i.e. the existence of the relevant legal rule, to be a reason which tips the balance and provides a sufficient reason for the required act. The way to interpret the fact that conformity is required even in the absence of other reasons for it is that the law itself is presented as such a reason. It does not matter if compliance is motivated by acknowledgement of such a claim. What matters is the nature of the claim itself.

The law's claim to legitimate authority is not merely a claim that legal rules are reasons. It includes the claim that they are exclusionary reasons for disregarding reasons for non-conformity. By and large I shall disregard this aspect of the law's claimed authority in the following essays. A few remarks may, however, be in place here. Many legal systems include doctrines specifically designed to allow individuals to deviate from an otherwise binding legal requirement if they consider themselves bound by strong reasons to do so. The rights of freedom of conscience and of conscientious objection are but prominent

<sup>2</sup> The legitimacy of so doing is explained in Essay 3 and other essays of Part II.

examples of such doctrines. In many countries prosecutorial discretion is meant by law to take account of such considerations. But in all legal systems the recognition given by law to such non-legal reasons is limited and is itself strictly regulated by law: non-legal reasons do not justify deviation from a legal requirement except if such justification is allowed by a specific legal doctrine.

It may be thought that the law ought to give recognition to all relevant considerations, that failure to do so is a moral defect in the law. In a sense it is a moral defect that the law sometimes does not enable courts to recognize the relevance of certain considerations. But if so, it is an inevitable defect. Perhaps many of the considerations not currently recognized by, for example, English courts should be recognized by them. But it is practically impossible for the law to recognize all the considerations relevant to cases to which it applies. If the law has moral authority then this authority must be established by showing, among other things, that it is better not to enable courts to apply certain considerations than not to have law at all.

The practical inability of the law to give full weight to all the reasons which may affect an agent is particularly obvious when the costs of enforcing proper recognition of such reasons or of the likely abuse of such a law if not adequately enforced seem to outweigh the benefits. This is likely to happen on numerous occasions. When concentrating attention on individual cases it is often felt that the arguments for making the case an exception to the rule are overwhelming and individuals or at the very least the courts must be given powers to recognize such exceptions, i.e. to have a general power to exempt from the law on grounds of hardship or justice. This may be justified with respect to certain legal problems. But in many areas this cure will be worse than the disease. It may be better to cause hardship in a few cases than to lead to great uncertainty in many. Uncertainty frustrates expectations, prevents planning, encourages abuses of the law by encouraging attempts to exploit its vague margins, and very often policing such laws involves prying into the private affairs of individuals. In matters where these considerations are important and where serious hardship or injustice is rare, certainty is at a premium and the courts are justifiably denied exemptive powers. The result is that very weighty reasons for deviating from the legally required act

on a specific occasion may have been excluded. Another cause for non-recognition is oversight. In drafting the law it may not have been realized that very pressing reasons for deviating from it exist or will exist in the future in a certain class of cases to which it applies. The slowness of the reforming and legislative processes often prevents law from being amended even after many are aware of its shortcomings.

I have mentioned practical inability and oversight as reasons for the law's failure to give full weight to all the reasons which may affect a case. These observations are made to explain not to justify. It is not here claimed that such reasoning always or indeed ever justifies the authority of the law. I shall return to this question in Part IV below.<sup>3</sup> The purpose of these remarks was merely to outline the kinds of argument normally invoked to explain and defend the fact that the law sometimes disregards valid and weighty reasons, for the problem is not how to justify these facts but how to interpret them.

The considerations mentioned are typical exclusionary reasons. Suppose there is a law requiring one to  $\phi$ . Let us assume that there is a very pressing reason,  $R$ , for not- $\phi$ -ing on a particular occasion. The law requires one not to act on that reason. Is it because it claims absolute weight for itself which defeats  $R$  or is it because the law claims to itself the status of a protected reason which excludes  $R$ ? One should look for the standard reasons for which  $R$  is rejected by the law. Those are the reasons expressed by the courts as justification for not excusing someone who acts on  $R$  and breaks the law and on similar occasions. My claim is that these reasons are those I mentioned above and other related reasons. Some will, therefore, be arguments that allowing everyone to act for  $R$  will have, on the whole, worse consequences because of the rate of failure to do so properly (people will be often wrong in thinking that  $R$  exists in their situation, etc.).

But the fact that trying to act for a certain reason ( $R$ ) for  $\phi$ -ing is more likely, because of the rate of failure to do so properly, to do more harm than good is not a reason for not- $\phi$ -ing.

<sup>3</sup> There is more than one question of justification involved. One must distinguish between 'Is the law (or legislator) right to claim such authority in those circumstances?' and 'Is the individual morally bound to acknowledge that authority?' Sometimes an affirmative answer to the first question will be combined with a negative answer to the second.

It is a reason for not acting for *R*. Indirectly it weakens the case for  $\phi$ -ing, but directly it is an exclusionary reason: a reason for excluding *R*. Similarly, the fact that, though the rule that one should  $\phi$  should be changed (because all things considered one should not  $\phi$ ) it should not be changed here and now or by me, is not a reason for  $\phi$ -ing which overrides the reasons for not- $\phi$ -ing but rather a reason why I should disregard those reasons for not- $\phi$ -ing. It is the essential exclusionary character of a rule that it resists permanent revision. It is immune from the claim that it should be re-examined with a view to possible revision on every occasion to which it applies.<sup>4</sup> This regulation of the appropriate method of revising rules is based not on the belief that until revised the reasons for the act required outweigh all conflicting reasons, but on the more rational and less mysterious belief that until it is revised it is justifiable to exclude such conflicting reasons.

One last caveat: what is excluded by a rule of law is not all other reasons, but merely all those other reasons which are themselves not legally recognized. We should not conceive of the law as a set of isolated norms each having its own separate and independent function, but as a set of (potentially conflicting or reinforcing) reasons which together determine what is required by law.

If this analysis is correct then the law claims authority. The law presents itself as a body of authoritative standards and requires all those to whom they apply to acknowledge their authority. This is of course not a very novel thesis. In one way or another it is common to think that the law claims authority. The purpose of this essay and of the previous one was to interpret that common notion, give it precise meaning, and vindicate it by integrating it to a wider thesis about the nature of authority and other known facts about the law. By doing this the central problem of this book was defined and presented. Since the law claims authority should its claim be acknowledged? Is it justified? Is there an obligation to submit to the authority of the law; if not, is it at all permissible to do so? Is the legitimacy of the authority of the law conditional on a legal recognition of a right of dissent?

<sup>4</sup> Cf. *Practical Reason and Norms* (2nd ed., Oxford, 1999), ch. 2.



PART II  
THE NATURE OF LAW AND  
NATURAL LAW



## Legal Positivism and the Sources of Law\*

### I. The Nature of Legal Positivism

The perennial and inexhaustible nature of the controversy concerning the positivist analysis of the law is due in no small measure to the elusive meaning of 'positivism' in legal philosophy. True, it is well established that legal positivism is essentially independent (even though not historically unrelated) both of the positivism of nineteenth-century philosophy and of the logical positivism of the present century. But the great variation between different positivist theories of law and the large variety of philosophical motivations permeating the work of the non-positivists indicate the difficulty, perhaps the impossibility, of identifying legal positivism at its source—in a fundamental positivist philosophical outlook. The easiest approach to the continuing controversy concerning legal positivism is through the particular theses or groups of theses round which it revolves.

Three areas of dispute have been at the centre of the controversy: the identification of the law, its moral value, and the meaning of its key terms. We could identify these as the social thesis, the moral thesis, and the semantic thesis respectively. It should be understood, however, that in each area positivists (and their opponents) are identified by supporting (or rejecting) one or more of a whole group of related theses rather than any particular thesis.

In the most general terms the positivist social thesis is that what is law and what is not is a matter of social fact (that is, the variety of social theses supported by positivists are various refinements and elaborations of this crude formulation). Their moral thesis is that the moral value of law (both of a particular law and of a whole legal system) or the moral merit it has is a contingent matter dependent on the content of the law and the circumstances of the society to which it applies. The only

\* My thinking on the problems discussed in this essay was greatly influenced by conversations with R. M. Dworkin and J. M. Finnis, who disagree with many of my conclusions.

semantic thesis which can be identified as common to most positivist theories is a negative one, namely, that terms like 'rights' and 'duties' cannot be used in the same meaning in legal and moral contexts. This vague formulation is meant to cover such diverse views as: (1) 'moral rights' and 'moral duties' are meaningless or self-contradictory expressions, or (2) 'rights' and 'duties' have an evaluative and a non-evaluative meaning and they are used in moral contexts in their evaluative meaning whereas in legal contexts they are used in their non-evaluative meaning, or (3) the meaning of 'legal rights and duties' is not a function of the meaning of its component terms—as well as a whole variety of related semantic theses.

Of these the social thesis is the more fundamental. It is also responsible for the name 'positivism' which indicates the view that the law is posited, is made law by the activities of human beings. The moral and semantic theses are often thought to be necessitated by the social thesis. In crude outline the arguments run as follows: Since by the social thesis what is law is a matter of social fact, and the identification of law involves no moral argument, it follows that conformity to moral values or ideals is in no way a condition for anything being a law or legally binding. Hence, the law's conformity to moral values and ideals is not necessary. It is contingent on the particular circumstances of its creation or application. Therefore, as the moral thesis has it, the moral merit of the law depends on contingent factors. There can be no argument that of necessity the law has moral merit. From this and from the fact that terms like 'rights' and 'duties' are used to describe the law—any law regardless of its moral merit—the semantic thesis seems to follow. If such terms are used to claim the existence of legal rights and duties which may and sometimes do contradict moral rights and duties, these terms cannot be used in the same meaning in both contexts.

I have argued elsewhere<sup>1</sup> that both arguments are fallacious and that neither the moral nor the semantic theses follow from the social one. The claim that what is law and what is not is purely a matter of social fact still leaves it an open question whether or not those social facts by which we identify the law or determine its existence do or do not endow it with moral

<sup>1</sup> *Practical Reason and Norms* (2nd ed., Oxford, 1999), pp. 162 ff., and see also Essay 8 below.

merit. If they do, it has of necessity a moral character. But even if they do not, it is still an open question whether, given human nature and the general conditions of human existence, every legal system which is in fact the effective law of some society does of necessity conform to some moral values and ideals. As for the semantic thesis, all the positivist has reason to maintain is that the use of normative language to describe the law does not always carry the implication that the speaker endorses the law described as morally binding. Put somewhat more precisely this means that normative language when used to state the law does not always carry its full normative force. To this even the non-positivist can agree. This does not justify the view that terms like rights and duties are used with a different meaning in legal and moral contexts.

It is not the purpose of the present essay to explore these arguments. I mention them only to indicate the extent to which the version of positivism that will be argued for here is a moderate one which need not conflict with the natural lawyer's view concerning the semantic analysis of normative terms and the relation between law and morality. The following are but some examples of views usually associated with natural law theories which are compatible with the version of positivism defended below:

- (a) 'A legal duty' means a duty which one has because the law requires the performance of that action.
- (b) There is a necessary connection between law and popular morality (i.e. the morality endorsed and practised by the population).
- (c) Every legal system's claim to authority is justified.

Whether or not these views are true, they are certainly compatible with the social thesis which is the backbone of the version of positivism I would like to defend. The social thesis is best viewed not as a 'first-order' thesis but as a constraint on what kind of theory of law is an acceptable theory—more specifically it is a thesis about some general properties of any acceptable test for the existence and identity of legal systems.

### **A. The (Strong) Social Thesis**

A jurisprudential theory is acceptable only if its tests for identifying the content of the law and determining its existence depend exclusively on facts of human

behaviour capable of being described in value-neutral terms, and applied without resort to moral argument.

This formulation is less clear than it might be. A more clear and lucid statement requires a fuller theoretical elaboration and is likely, therefore, to be more controversial. The above formulation strives to get at the core motivation and the basic idea underlying the various formulations of the social thesis and accepts the inevitable cost in lack of precision. Some clarification may nevertheless be called for.

First, the thesis assumes that any complete theory of law includes tests for the identification of the content and determination of the existence of the law. This seemed self-evident to many philosophers of law who saw it as one of their main tasks to provide such tests. Other equally influential legal philosophers were never stirred to do so and felt that such tests are no part of legal philosophy or are fruitless or impossible. Lon Fuller is the most eminent of those contemporary philosophers who have taken such a view. The reasons for dissenting from such positions will be indicated briefly in the next section. It is best to regard such theories as incomplete theories of law. For one reason or another most, if not all, theories of law are incomplete in that they do not propose answers to some questions which fall within the province of jurisprudence.

Secondly, the thesis assumes that there is a sufficiently rich vocabulary of value-neutral terms. It does not assume that there is a clear and sharp break between value-laden and value-free terms. Nor is it committed to any side in the naturalist/anti-naturalist dispute. That the test is *capable* of being described in value-neutral terms does not mean that no value or deontic conclusions are entailed by it. To assert that is to take an anti-naturalist position.

Thirdly, the thesis does not require disregarding the intentions, motivations, and moral views of people. Value-neutrality does not commit one to behaviourism.

Finally, it is worth noting that the social thesis can be divided into two: *A*—A social condition is necessary for identifying the existence and content of the law: A rule is a legal rule only if it meets a social condition. *B*—A social condition is sufficient for identifying the existence and content of the law: A rule is a legal rule if it meets the social condition.

## II. The Social Thesis

I have claimed that the social thesis has always been at the foundation of positivist thinking about the law and that its semantic and moral consequences have all too often been misunderstood. It is not to my purpose here to expound and defend any particular view about the tests by which the existence and content of law are to be identified.<sup>2</sup> But since acceptance of the social thesis does give shape to theories of law which endorse it, it is important to reflect once again upon the reasons supporting the social thesis. In so doing I will inevitably commit myself to certain more definite views about the social conditions for the existence and identity of legal systems.

A. The most general and non-theoretical justification of the social thesis is that it correctly reflects the meaning of 'law' and cognate terms in ordinary language. This claim can be and has been illustrated often enough. It seems fundamentally sound as an essential part of every defence of the social thesis and yet in itself it is inconclusive. The word 'law' has non-legal uses: laws of nature, moral laws, laws of various institutions, the laws of thought, etc. Several of these have problematical status; moreover there are no clear demarcation lines in linguistic usage between the different kinds of law. Hence the dispute about the character of international law, for example, cannot be determined by appeal to ordinary language.

For similar reasons usage is too amorphous to give adequate support for the social thesis. It certainly suggests that law has a social base, that Nazi Germany had a legal system, etc. But it is not sufficiently determinate to establish beyond dispute that social facts are both sufficient and necessary conditions for the existence and identity of the law.

Finally, we do not want to be slaves of words. Our aim is to understand society and its institutions. We must face the question: is the ordinary sense of 'law' such that it helps identify facts of importance to our understanding of society?

B. The social thesis is often recommended on the ground that it clearly separates the description of the law from its evaluation. This, it is alleged, prevents confusion and serves clarity

<sup>2</sup> See Essays 5 and 6 below and my *Practical Reason and Norms*, sections 4.3–5.2, where I suggested various modifications and elaborations on Hart's ideas as expounded in *The Concept of Law* (Oxford, 1961).

of thought. This is true, but it presupposes the thesis rather than supports it. If the law is to be identified by social tests then trying to identify it without clearly separating social facts from evaluative considerations is misleading and often downright wrong. But if the identification of the law involves, as many natural lawyers believe it does, evaluative as well as social conditions then to distinguish between the two in identifying the law is misleading and wrong.

C. Adhering to the social thesis eliminates investigator's bias. It requires that the investigator should put aside his evaluative and deontic views and rely exclusively on considerations which can be investigated and described in a value-neutral way. This again, though true, presupposes the social thesis and is one of its results rather than its foundation. For in this respect too it must be admitted that if those natural lawyers who reject the social thesis are right then involving the investigator's sense of values (it will not then be called bias) is the only proper way for identifying the law. This does not mean that on this view the law is what it is because the investigator believes in certain values. It does, however, mean that the proper way for identifying the law is to inquire into the validity and implications of certain values.

D. There are, no doubt, many other reasons and variations on reasons which have been proposed in support of the social thesis and many of them have at least some truth in them. But the main justification of the social thesis lies in the character of law as a social institution. Some social institutions may have to be understood in ways which are incompatible with an analogous social thesis applied to them. But the law, like several others, is an institution conforming to the social thesis. To see this, it is necessary to specify in a general way the main ingredients of the tests for existence and identity for a legal system and to identify those with which the social thesis is concerned. The tests for identity and existence of a legal system contain three basic elements: efficacy, institutional character, and sources.

Efficacy is the least controversial of these conditions. Oddly enough it is also the least studied and least understood. Perhaps there is not much which legal philosophy can contribute in this respect. Though I believe there are at least some, however

elementary, difficulties which need to be explored.<sup>3</sup> Since this essay is not concerned with the precise details of the efficacy condition these difficulties can be overlooked. Suffice it that all agree that a legal system is not the law in force in a certain community unless it is generally adhered to and is accepted or internalized by at least certain sections of the population. This condition is simply designed to assure that the law referred to is the actual law of a given society and not a defunct system or an aspiring one. It is the least important of the conditions. It is not disputed by natural lawyers. And it does not help to characterize the essence of law as a kind of human institution. It distinguishes between effective and non-effective law and not between legal and non-legal systems. Consider, by way of analogy, social morality. The same condition applies. No morality is the social morality of a population unless it is generally conformed to and accepted by that population. Here the condition of efficacy does not illuminate the nature of morality. It merely tells an effective morality from one which is not.

More important and also more controversial is the second component of the tests for existence and identity—the institutionalized character of the law. Again, the many controversies about the precise nature of the institutional aspect of law can be side-stepped here. It is widely agreed (and by many natural lawyers as well) that a system of norms is not a legal system unless it sets up adjudicative institutions charged with regulating disputes arising out of the application of the norms of the system. It is also generally agreed that such a normative system is a legal system only if it claims to be authoritative and to occupy a position of supremacy within society, i.e. it claims the right to legitimize or outlaw all other social institutions.

These institutionalized aspects of law identify its character as a social type, as a kind of social institution. Put in a nutshell, it is a system of guidance and adjudication claiming supreme authority within a certain society and therefore, where efficacious, also enjoying such effective authority. One may think that there is much more that can be said about the sort of social institution that law is. Why be so sparing and abstract in its description? No doubt the features of law mentioned can and

<sup>3</sup> See my *The Concept of a Legal System* (Oxford, 1970), ch. 9, for some of the puzzling aspects of many common views about efficacy.

should be elaborated in much greater detail. But when articulating a general test for existence and identity for the law one probably should not go beyond this bare characterization. The rest belongs properly to the sociology of law, for it characterizes some specific legal systems or some types of legal systems (modern capitalist, feudal, etc.), and not necessarily all legal systems.<sup>4</sup>

'Law', as was mentioned already, is used in many different contexts and is applied to norms of great variety and diversity. Lawyers quite naturally focus their professional attention on a certain range of uses: those which are tied to institutions of the type described. Many legal philosophers have suggested that the philosophical analysis of law should follow the legal profession and should pin its analysis to this kind of institution. This is quite natural and completely justified. Given even the very sketchy and rudimentary characterization proposed above, it is amply clear that law thus understood is an institution of great importance to all those who live in societies governed by law, which nowadays means almost everybody. There is more than enough justification to make it a subject of special study (which need not and should not neglect its complex interrelations with other institutions and social forces). There is also sufficient reason to encourage the general public's consciousness of law as a special type of institution.

Many natural law theories are compatible with all that was said above concerning the institutional nature of law. Yet it must be pointed out that such an institutionalized conception of law is incompatible with certain natural law positions; and this for two reasons. In the first place, it is a consequence of the institutionalized character of the law that it has limits. Legal systems contain only those standards which are connected in certain ways with the operation of the relevant adjudicative institutions.<sup>5</sup> This is what its institutionalized character means.

<sup>4</sup> The main possible addition to the facts I mentioned are sanctions, the use of coercion or of force, and the existence of institutions for law-enforcement. See on this H. Oberdiek, 'The Role of Sanctions and Coercion in Understanding Law and Legal Systems', *American Journal of Jurisprudence* (1975), 71; Raz, *Practical Reason and Norms*, pp. 154-62.

<sup>5</sup> Kelsen thought the relation is simple: Laws are norms addressed to courts (see, for example, *The General Theory of Law and State* (New York, 1945), p. 29). Others suggest more indirect connections. Most notable is Hart's idea that laws are standards courts are bound to apply and use in adjudication: *The Concept of Law*, pp. 89 ff.

Hence the law has limits: it does not contain all the justifiable standards (moral or other) nor does it necessarily comprise all social rules and conventions. It comprises only a subset of these, only those standards having the proper institutional connection.<sup>6</sup> This is incompatible with the view that law does not form a separate system of standards and especially with the claim that there is no difference between law and morality or between it and social morality.

A second and perhaps more radical consequence of the conception of law as an institutional system is that one cannot impose moral qualifications as conditions for a system or a rule counting as legal which are not reflected also in its institutional features. If law is a social institution of a certain type, then all the rules which belong to the social type are legal rules, however morally objectionable they may be. Law may have necessary moral properties, but if so, then only on the ground that all or some of the rules having the required institutional connections necessarily have moral properties. To impose independent moral conditions on the identity of law will inevitably mean either that not all the rules forming a part of the social institution of the relevant type are law or that some rules which are not part of such institutions are law. Either way 'law' will no longer designate a social institution.

### III. The Sources of Law

Most positivists are ambiguous concerning one interesting point. While their general terms suggest an endorsement of the strong social thesis, their actual doctrines rest on efficacy and institutionality as the only conditions concerning the social foundation of the law. Let the combination of these two conditions be called the weak social thesis. It is easy to show that the weak and the strong theses are not equivalent. Suppose that the law requires that unregulated disputes (i.e. those with

<sup>6</sup> Cannot a society have judicial institutions instructed to apply all social rules and cannot we envisage such a society lacking a clear differentiation between social morality and ideal or critical morality? Such societies are possible and probably have existed. This, however, merely shows that from their point of view there was no distinction between law and morality (unless they were made conscious of the distinction by observing other communities). We who have the distinction can still apply it to them when judging that, as things stand in their community, law encompasses the whole of their social morality. But things could have been different even for them.

respect to which the law is unsettled) be determined on the basis of moral considerations<sup>7</sup> (or a certain subclass of them, such as considerations of justice or moral considerations not fundamentally at odds with social morality). Suppose further that it is argued that in virtue of this law moral considerations have become part of the law of the land (and hence the law is never unsettled unless morality is). This contention runs directly counter to the strong thesis. If it is accepted, the determination of what is the law in certain cases turns on moral considerations, since one has to resort to moral arguments to identify the law. To conform to the strong thesis we will have to say that while the rule referring to morality is indeed law (it is determined by its sources) the morality to which it refers is not thereby incorporated into law. The rule is analogous to a 'conflict of law' rule imposing a duty to apply a foreign system which remains independent of and outside the municipal law.

While all this is clear enough, it is equally clear that the contrary view (according to which morality becomes part of the law as a consequence of the referring law) does not offend against the requirement of efficacy. For here too the bulk of the legal system may be conformed to. Nor is this view inconsistent with the institutional aspect of law: morality becomes law, on this view, by being tied to the relevant institutions. Finally, the allegation that morality can be thus incorporated into law is consistent with the thesis of the limits of law, for it merely asserts that source-based laws may from time to time incorporate parts of morality into law while imposing perhaps various conditions on their applicability. Having said that I should add that the result of admitting the view under consideration is that some non-source-based moral principles are part of almost every legal system, since most legal systems require judges to apply moral considerations on various occasions.

The difference between the weak and the strong social theses is that the strong one insists, whereas the weak one does not, that the existence and content of every law is fully determined by social sources. On the other hand, the weak thesis, but not the strong one, builds into the law the conditions of efficacy and

<sup>7</sup> Note that the reference is to morality, not to social morality. Social morality is based on sources: the customs, habits, and common views of a community.

institutionality. The two theses are logically independent. The weak thesis though true is insufficient to characterize legal positivism. It is compatible with:

(a) Sometimes the identification of some laws turns on moral arguments,

but also with

(b) In all legal systems the identification of some laws turns on moral argument.

The first view is on the borderline of positivism and may or may not be thought consistent with it. But whereas the first view depends on the contingent existence of source-based law making moral considerations into the criteria of validity in certain cases (as in the example above), the second view asserts a conceptual necessity of testing law by moral argument and is clearly on the natural law side of the historical positivist/natural law divide.

I will argue for the truth of the strong social thesis (thus excluding both (a) and (b)).<sup>8</sup> I shall rename the strong social thesis 'the sources thesis'. A 'source' is here used in a somewhat technical sense (which is, however, clearly related to traditional writings on legal sources). A law has a source if its contents and existence can be determined without using moral arguments (but allowing for arguments about people's moral views and intentions, which are necessary for interpretation, for example). The sources of a law are those facts by virtue of which it is valid

<sup>8</sup> The weak social thesis provides all the ingredients by which one determines whether a normative system is a legal system and whether it is in force in a certain country. In other words, the weak social thesis provides a complete test of existence of legal systems, a test by which one determines whether there is a legal system in force in a country. It also contributes (i.e. the institutional character of law contributes) some of the ingredients which make up the test of identity of a legal system (i.e. the test by which one determines whether two norms belong to the same legal system), but here it is insufficient and has to be supplemented by the strong social thesis, i.e. by the claim that all laws have social sources.

E. P. Soper, 'Legal Theory and the Obligation of the Judge: The Hart/Dworkin Dispute', *Michigan L. Rev.* 75 (1977) pp. 511 f., and D. Lyons, 'Principles, Positivism and Legal Theory', *Yale L.J.* 87 (1977) pp. 424 f., argue that legal positivism is consistent with (a). Supporters of such a conception of the law have to provide an adequate criterion for separating legal references to morality, which make its application a case of applying pre-existing legal rules, from cases of judicial discretion in which the judge, by resorting to moral considerations, is changing the law. I am not aware of any serious attempt to provide such a test.

and which identify its content. This sense of 'source' is wider than that of 'formal sources' which are those establishing the validity of a law (one or more Acts of Parliament together with one or more precedents may be the formal source of one rule of law). 'Source' as used here includes also 'interpretative sources', namely all the relevant interpretative materials. The sources of a law thus understood are never a single act (of legislation, etc.) alone, but a whole range of facts of a variety of kinds.

What are the reasons for accepting the sources thesis? Two arguments combine to support it. The one shows that the thesis reflects and explicates our conception of the law; the second shows that there are sound reasons for adhering to that conception.

When discussing appointments to the Bench, we distinguish different kinds of desirable characteristics judges should possess. We value their knowledge of the law and their skills in interpreting laws and in arguing in ways showing their legal experience and expertise. We also value their wisdom and understanding of human nature, their moral sensibility, their enlightened approach, etc. There are many other characteristics which are valuable in judges. For present purposes these two kinds are the important ones. The point is that while it is generally admitted that both are very important for judges as judges, only the first group of characteristics mentioned is thought of as establishing the legal skills of the judge. The second group, though relevant to his role as a judge, is thought of as reflecting his moral character, not his legal ability. Similarly, when evaluating judgments as good or bad, lawyers and informed laymen are used to distinguishing between assessing judicial arguments as legally acceptable or unacceptable and assessing them as morally good or bad. Of many legal decisions we hear that they are legally defective, being based on a misinterpretation of a statute or a case, etc. Of others it is said that though legally the decisions are acceptable, they betray gross insensitivity to current social conditions, show how conservative judges are, that they are against trade unions, or that in their zeal to protect individuals they go too far in sacrificing administrative efficiency, etc.

These distinctions presuppose that judges are, at least on

occasion, called upon to rely on arguments revealing their moral character rather than their legal ability. (It is unreasonable to suppose that the judge's moral character reveals itself only when he is wrong in law. It affects decisions too often for that to be a reasonable hypothesis.) As indicated above, the use of moral judgment is regarded not as a special case of applying law or legal arguments, but is contrasted with them. This is manifested in the way the two kinds of tests evaluating judges and judgments are related to two further distinctions. The first is that between applying the law and creating, innovating, or developing the law. It is a common view that judges both apply the law and develop it. And though their two functions are extremely hard to disentangle in many cases, yet sometimes, at least, it is clear of a case that it breaks new ground, while of many others it may be equally clear that they merely apply established law. The important point is that it is our normal view that judges use moral arguments (though perhaps not only such arguments) when developing the law and that they use legal skills when applying the law (though not only legal skills are used when they have to decide whether to apply a precedent or distinguish or overrule it. I shall disregard this problem in the sequel and will return to it in Essay 10 below).

Finally there is the distinction between settled and unsettled law. All lawyers know that on some questions the law is unsettled. Sometimes they say on such cases that no one knows what the law is—as if there is law on the question which is very difficult to discover. But most of the time they express themselves more accurately, saying that this is an open question, that the law is unsettled, etc. (On the interpretation of these expressions, see Essay 4.) It is primarily in deciding cases regarding which the law is unsettled (as well as in distinguishing and reversing settled law) that judges are thought to develop the law using moral, social, and other non-legal arguments. It is when deciding cases where the law is settled that the judges are thought of as using their legal skills in applying the law.

The sources thesis explains and systemizes these distinctions. According to it, the law on a question is settled when legally binding sources provide its solution. In such cases judges are typically said to apply the law, and since it is source-based, its application involves technical, legal skills in reasoning from

those sources and does not call for moral acumen. If a legal question is not answered by standards deriving from legal sources then it lacks a legal answer—the law on the question is unsettled. In deciding such cases courts inevitably break new (legal) ground and their decision develops the law (at least in precedent-based legal systems). Naturally, their decisions in such cases rely at least partly on moral and other extra-legal considerations.

One need not assume complete convergence between the distinctions mentioned above and the sources thesis. If, in fact, the sources thesis coincides with the way these distinctions are generally applied, it has explanatory power and is supported to that extent. It can then be regarded as being a systemizing or a tidying-up thesis where it goes beyond the ordinary use of these distinctions. This argument for the sources thesis is not an argument from the ordinary sense of ‘law’ or any other term. It relies on fundamental features of our understanding of a certain social institution, the primary examples of which are contemporary municipal legal systems but which extend far beyond them. It is not part of the argument that a similar conception of legal systems is to be found in all cultures and in all periods. It is part of our ways of conceiving and understanding the working of social institutions. There is nothing wrong in interpreting the institutions of other societies in terms of our typologies. This is an inevitable part of any intelligent attempt to understand other cultures. It does not imply that in interpreting alien institutions you disregard the intentions, beliefs, or value-schemes of their participants. It only means that at some stage you classify their activities, thus interpreted, in terms of a scheme for analysing social institutions of which the participants themselves may have been ignorant.

Still, it may be reassuring to know that the sources thesis is not merely a reflection of a superficial feature of our culture. I shall argue briefly that the sources thesis captures and highlights a fundamental insight into the function of law. It is a commonplace that social life requires and is facilitated by various patterns of forbearances, co-operation, and co-ordination between members of the society or some of them. The same is true of the pursuits of goals which the society or sections in it may set themselves. Different members and different sections

of a society may have different views as to which schemes of co-operation, co-ordination, or forbearance are appropriate. It is an essential part of the function of law in society to mark the point at which a private view of members of the society, or of influential sections or powerful groups in it, ceases to be their private view and becomes (i.e. lays a claim to be) a view binding on all members notwithstanding their disagreement with it. It does so and can only do so by providing publicly ascertainable ways of guiding behaviour and regulating aspects of social life.<sup>9</sup> Law is a public measure by which one can measure one's own as well as other people's behaviour. It helps to secure social co-operation not only through its sanctions providing motivation for conformity but also through designating in an accessible way the patterns of behaviour required for such co-operation. This fact has been emphasized by many a natural lawyer for it forms part of the justification of the need for positive law. Locke is a prominent and well-known example. Hart more than anybody else emphasized the point among legal positivists.

To prevent misunderstanding let me elaborate some of the crucial steps in the argument. Many societies (large or small) have a relatively formal way of distinguishing between expressions of views, demands, etc., and authoritative rulings. Such a distinction is an essential element in our conception of government, be it in a family, in a loosely organized community, or in the state. Not all authoritative rulings are laws, not all systems of such rules are legal systems. But marking a rule as legally binding is marking it as an authoritative ruling. This marking-off of authoritative rulings indicates the existence in that society of an institution or organization claiming authority over members of the society that is holding them bound to conform to certain standards just because they were singled out by that purported authority regardless of whether or not they are justifiable standards on other grounds. Since it is of the very essence of the alleged authority that it issues rulings which are binding regardless of any other justification, it follows that it

<sup>9</sup> I do not mean to suggest that all laws are open. Secret laws are possible provided they are not altogether secret. Someone must know their content some of the time. They are publicly ascertainable and they guide the behaviour of the officials to whom they are addressed or who are charged with their enforcement by being so.

must be possible to identify those rulings without engaging in a justificatory argument, i.e. as issuing from certain activities and interpreted in the light of publicly ascertainable standards not involving moral argument.

If the first argument for the sources thesis was that it reflects and systemizes several interconnected distinctions embedded in our conception of the law, the second argument probes deeper and shows that the distinctions and the sources thesis which explicates them help to identify a basic underlying function of the law: to provide publicly ascertainable standards by which members of the society are held to be bound so that they cannot excuse non-conformity by challenging the justification of the standard. (Though, of course, in many countries they are free to act to change it.) This is the reason for which we differentiate between the courts' applying the law, i.e. those standards which are publicly ascertainable and binding beyond a moral argument open to the litigants, and the activity of the courts in developing the law relying on moral and other rational considerations. In making this a test for the identification of law, the sources thesis identifies it as an example of a kind of human institution which is of decisive importance to the regulation of social life.

## Legal Reasons, Sources, and Gaps\*

It is common ground to all legal positivists that the law has social sources, i.e. that the content and existence of the law can be determined by reference to social facts and without relying on moral considerations. This view led early positivists such as Bentham, Austin, and to a degree also the American Realists to a reductivist interpretation of legal statements. They claimed that legal statements are synonymous with statements about what certain people commanded or willed, or about the chances that a man may come to harm of a certain kind, or about the likelihood that courts will reach certain decisions. Professor Hart, while accepting the sources thesis, mounted a most formidable criticism of reductivism. He argued that legal statements are deontic or practical. They are used to demand and justify action and thus function in discourse and argument in ways which no theoretical statements could.<sup>1</sup>

Professor Hart's arguments are convincing and there is no need to elaborate them here. His own several attempts to reconcile the sources thesis, according to which all laws have a source, with his rejection of reductivist explanations of legal statements are less successful.<sup>2</sup> More illuminating in this respect is Kelsen's work: Kelsen never succumbed to the reductivist temptation but developed a most ingenious way of combining the sources thesis with anti-reductivism.<sup>3</sup> Unfortunately, his discussion is so obscure that he is very often mistaken for a reductivist.

\* Published in *Archiv für Rechts- und Sozialphilosophie, Beiheft 11* (1979).

<sup>1</sup> For Hart's repeated attempts to come to grips with the problem of reductivism see especially 'The Ascription of Responsibility and Rights', *Proceedings of the Aristotelian Society* 49 (1948-9), 171; *Definition and Theory in Jurisprudence* (Oxford, 1953); and *The Concept of Law* (Oxford, 1961).

<sup>2</sup> For a recent penetrating analysis of the development of Hart's views on the issue and a powerful criticism of his views see G. P. Baker, 'Defeasibility and Meaning' in P. M. S. Hacker and J. Raz (eds.), *Law, Morality and Society* (Oxford, 1977).

<sup>3</sup> His views remain the same in *General Theory of Law and State* (New York, 1944) and *Pure Theory of Law* (2nd ed., Berkeley, 1967).

Kelsen's theory, though containing invaluable insights, is defective on various counts.<sup>4</sup> The purpose of the following remarks is to contribute<sup>5</sup> to an anti-reductivist explanation of legal statements, based on the sources thesis. Such an explanation, it will be argued, presupposes the possibility of gaps in the law and a proper understanding of the relation between law and legal adjudication.

## I. Professor Dworkin's Refutation of Positivism

If the law is based on sources, what is the law on matters not referred to in any source? If a person has a legal right only if its existence can be traced back to a legal source, what is his situation if no source confers it on him? Is it then the law that he does not have it? Is that the case even if there is no law (traceable to a source) denying him the right? Are there gaps in the law? What does it mean to say that there is a legal gap in a situation in which the judge is instructed how to find a solution by the law (e.g. 'Act as if you are a legislator')?

These and similar questions are the subject of many jurisprudential writings.<sup>6</sup> Professor Dworkin has attempted to show that endorsement of the sources thesis makes these questions insoluble and that therefore legal positivism is an incoherent doctrine.<sup>7</sup> His argument can serve as an introduction to an appreciation of the problem facing the supporter of the sources thesis.

### A. The Refutation

Professor Dworkin's argument is lengthy and designed to allow for various ways of interpreting the sources thesis. The following version of it seems to capture the main thrust of his attack: Assume that  $p$  is a variable ranging over legal propositions and that  $S(p)$  asserts the existence of a

<sup>4</sup> I have argued against reductivist interpretations of Kelsen in Essay 7 below and in 'Kelsen's General Theory of Norms', *Philosophia* (1976), where his views are also criticized.

<sup>5</sup> For other pieces of the jigsaw cf. 'Legal Validity' below.

<sup>6</sup> Cf. Alchouron and Bulygin, *Normative Systems* (Vienna, 1971), for an excellent discussion from which I have greatly benefited. My conclusions differ from theirs mainly because they do not allow for normative conflicts (which do not amount to inconsistencies) or for truth-value gaps. Nor do they draw the conclusions that I do from the sources thesis.

<sup>7</sup> See his 'No Right Answer?' in Hacker and Raz (eds.), *Law, Morality and Society*, some of the arguments of which he repeats in 'Can Rights be Controversial?', *Taking Rights Seriously* (London, 1977).

source for  $p$ . Assume a reductivist version of positivism by which  $p$  means the same as  $S(p)$ . It follows that:

$$(1) \vdash p \leftrightarrow S(p).$$

This entails both

$$(2) \vdash \neg p \leftrightarrow \neg S(p) \quad (\text{counterposition of (1)}).$$

$$(3) \vdash \neg p \leftrightarrow S(\neg p) \quad (\text{Substitution in (1)}).$$

Therefore (1) entails

$$(4) \vdash \neg S(p) \leftrightarrow S(\neg p).$$

That is, it is a theorem that whenever  $p$  has no source (e.g.  $p$  has not been enacted nor made the *ratio decidendi* of a binding decision, etc.) *not-p* has a source (e.g. it has been enacted that *not-p*). This conclusion is patently false. In England, for example, there is no source for the legal proposition that it is legally prohibited to kill any butterfly, but neither is there a source for its contradiction, i.e. that it is not legally prohibited to kill any butterfly. (1) must be rejected for it entails a false conclusion.

## B. Positive/Negative Solution

A legal positivist cannot accept an interpretation of the sources thesis which entails (1). What routes are available to him to escape Professor Dworkin's criticism? Three different solutions will be outlined briefly in this section. In the next section I shall argue in greater detail for a fourth. First, one may insist that different categories of legal propositions should be distinguished. (1) must be restricted to one category and not every legal proposition should be allowed to be substituted for  $p$  in it. In particular, positive legal propositions should be distinguished from negative ones. The sources thesis concerns positive propositions only and negative ones cannot, on this view, be substituted for  $p$  in (1). This would block the inference of (3) and therefore of (4) from (1). Let us assume that we follow such a line of reasoning and distinguish between positive legal propositions ( $Pp$  being a variable ranging over them only) and negative ones ( $Np$ ). A negative proposition is true if and only if the corresponding positive one is false and false if and only if the corresponding positive proposition is true.

Hence:

$$(5) \vdash Np \leftrightarrow \neg Pp.$$

The sources thesis is understood as entailing

$$(6) \vdash Pp \leftrightarrow S(Pp).$$

There is an additional theorem for negative propositions:

$$(7) \vdash S(Np) \rightarrow Np.$$

(5) and (6) entail

(8)  $\vdash \neg S(Pp) \rightarrow Np$  (counterposition of (6) and substitution of equivalents in it).

Thus there will be two consistent theorems for relating negative propositions and sources ((7) and (8)) and one for positive propositions (6).

This move would not be more than a trick unless positive and negative legal propositions can be distinguished in such a way that there is an independent reason for treating the resulting kinds of propositions differently. Such reasons, fortunately, have often been pointed to before. It is a feature of legal language that whereas one needs a law or some legal source to support a claim that a person is under a duty, is lawfully married, or that a contract or a will is valid, absence of a source for the contrary suffices to show that no duty, marriage, or contract, etc., exists. A valid contract must be made, the validity of marriage depends on certain formal steps being taken, a law is needed to impose a duty, etc. But there need be no making of uncontracts for there not to be a contract, no ceremony for two persons to remain unmarried to each other, and no law to confer a non-duty on a person.

Pointing to the obvious asymmetries between positive and negative legal propositions is not explaining them. Nor can it be denied that many cases create problems. Does one need a source to have a right, or does one have a right unless there is a source denying it? Such problems call for further clarifications which in themselves shed more light on the nature of the sources thesis.<sup>8</sup>

<sup>8</sup> Cf. my brief discussion in *The Concept of a Legal System* (Oxford, 1970), pp. 170–2. The discussion of this issue below is merely an elaboration of the points made there.

### C. The True/False Solution

The sources thesis can, but need not, be defended through reliance *on* the distinction between positive and negative statements. If one does not wish to rely on (6) and (7) with their presupposition of a distinction between negative and positive legal propositions, one can accept (9) and (10) instead:

- (9)  $S(p)$  is a logically necessary and sufficient condition for the truth of  $p$ .  
 (10)  $S(\neg p)$  is a logically necessary and sufficient condition for the falsity of  $p$ .

This solution is offered by Professor Dworkin himself. Since  $\neg(p \ \& \ \neg p)$  and since (9) entails that  $S(\neg p)$  is necessary and sufficient for the truth of  $\neg p$  this solution entails:

- (11)  $\vdash S(\neg p) \rightarrow \neg S(p)$  and therefore  
 (12)  $\vdash \neg(S(\neg p) \ \& \ S(p))$ .

(12) may seem counter-intuitive. It means that if  $\neg p$  has a source (say it was enacted) then  $p$  has no source (was not enacted nor made a binding precedent, etc.). Cannot there be a situation in which both  $p$  and *not-p* have been validly enacted? In such a case  $S(\neg p) \ \& \ S(p)$  is true, though its negation is entailed by (9) and (10). But all this objection amounts to is that sources have to be defined so that  $\neg S(\neg p)$  is part of the source of  $p$ . The possibility of a negative condition like this being part of a source is no more than the possibility of repeal by enacting the opposite of the previous law, for example, repealing a permission by enacting a prohibition or a prohibition by enacting a permission. Since a source sufficient to establish the truth of a legal proposition must guarantee that the law on which it is based has not been repealed, it must include such negative conditions ( $\neg S(\neg p)$ ) wherever this form of repeal is legally recognized. If one is to preserve the logical consistency of legal language then one must postulate that where positive law fails to provide a rule of precedence, i.e. a rule determining which of two conflicting enactments, etc., is ineffective (or which of them is repealed) then both are ineffective so that neither  $S(p)$  nor  $S(\neg p)$  is true.<sup>9</sup>

<sup>9</sup> The same stipulation must be made in the version of the positive/negative solution presented above. My own position as set out below does not require this kind of stipulation.

The true/false solution differs from the negative/positive solution in presupposing the possibility of legal gaps. In a sense even Dworkin's positivist does not rule out the possibility of legal gaps. But with him as well as with the positive/negative solution in the version given above, legal gaps depend on non-legal gaps. According to them, if it could be the case that neither  $S(p)$  nor  $\neg S(p)$ , then it could also be that neither  $p$  nor  $\neg p$ . This possibility is open under the true/false solution as well, but to it is added the possibility of gaps not deriving from non-legal gaps but arising whenever it is neither the case that  $S(p)$  nor that  $S(\neg p)$ . When this is the case, there is neither a source for  $p$  nor for  $\neg p$ . Since unlike the negative/positive solution, the true/false solution does not distinguish between positive and negative propositions and requires a source for either, it must and does concede the possibility that for some  $p$  neither  $p$  nor  $\text{not-}p$  have a source. In such a case neither  $p$  nor  $\text{not-}p$  will be either true or false.

Notice that such a gap does not prevent one from defining a permission as the negation of a prohibition. 'x is legally permitted to  $\phi$ ' means the same as 'it is the law that x is not prohibited to  $\phi$ ' ( $LPer\ x, \phi = df. L - Pr\ x, \phi$ ). But this should be distinguished from: it is not the law that x is prohibited to  $\phi$  ( $\neg LPr\ x, \phi$ ). According to the true/false solution it cannot be the case that one is legally permitted to do that which one is legally prohibited from doing:

$$(13) \vdash \neg(LPr\ x, \phi \ \& \ LPer\ x, \phi).$$

For by definition

$$(14) \vdash LPr\ x, \phi \leftrightarrow L - Per\ x, \phi.$$

But it is possible that an act is neither prohibited nor permitted:

$$(15) \vdash \diamond ((\neg LPr\ x, \phi) \ \& \ (\neg LPer\ x, \phi)).$$

When an action is neither legally prohibited nor legally permitted there is a legal gap. (15) is just a special case of the general case of legal gaps allowed by the true/false solution where neither  $p$  nor  $\neg p$  is true.

The true/false approach may be thought the most faithful to the sources thesis since it regards no legal statement as true unless it has a source. Dworkin's positivist and the positive/

negative solutions allow that negative legal propositions can be true even if they have no source other than the absence of a source for their contradictory proposition.

In 'No Right Answer?' Professor Dworkin denies that the true/false solution is part of an acceptable theory of law: 'Everything turns on the question of which ground-rules of assertion the participants in the legal practice in fact follow, and we can easily imagine ground-rules very different from those that positivism presupposes' (p. 74). By this he must mean that the courts of a certain country may follow a practice of respecting certain considerations rather than others on grounds other than their sources, i.e. grounds determined not by social facts but by moral values.<sup>10</sup> This is, of course, possible. The point of the sources thesis is not that courts never rely on sourceless considerations, but rather that when doing so they are not relying on legally binding considerations but exercising their own discretion. Professor Dworkin does not refute legal positivism. He prefers to ignore it. But his oversight is nevertheless helpful. It draws attention to the problem of the relation between law and judicial jurisdiction. If the sources thesis entails that there is jurisdiction despite gaps, it is bound to explain how such jurisdiction is exercised.

#### **D. The Prima Facie Solution**

The prima facie solution has not been explored by legal philosophers. It is based on the idea that binding, valid laws may conflict and that such conflicts are not inconsistencies. They show that a law has prima facie force only. What is conclusively required by law in any given situation cannot, therefore, be judged by consulting one law only but is often the result of the operation of various laws, some of which may well conflict. Consequently two kinds of legal propositions should be distinguished: propositions about what is prima facie the case and propositions about what is conclusively the case.<sup>11</sup>

<sup>10</sup> Professor Dworkin must mean that if his argument is even to appear to have any force against the sources thesis which he stated to be his target in this section of his article.

<sup>11</sup> Professor Dworkin's 'The Model of Rules I', *Taking Rights Seriously*, which was first published in *Chicago Law Review* (1967) did a great deal to draw the attention of legal philosophers to the prima facie character of many laws. Cf. also my 'Legal Principles and the Limits of Law', *Yale L. J.* 81 (1972), 823, for an argument that the prima facie character of law is more extensive and radical than Professor Dworkin allows.

The prima facie solution interprets the sources thesis as applying to prima facie legal propositions ( $Pf\ p$ , for short). (1) is therefore rejected in favour of:

$$(16) \vdash Pf\ p \leftrightarrow S(p).$$

It entails the equivalent of (2), namely

$$(17) \vdash -S(p) \leftrightarrow -Pf\ p.$$

But the equivalent of (3) becomes

$$(18) \vdash S(-p) \leftrightarrow Pf-p.$$

There is nothing which could count as a source of  $-Pf\ p$ . One can legislate that there is a prima facie duty to  $\phi$ , but one cannot legislate that there is no prima facie duty to  $\phi$  (as distinct from ‘there is a prima facie duty not to  $\phi$ ’). If there is no prima facie duty to  $\phi$ , this cannot be the result of a special source. It can only result from the absence of a source for the contrary (as in (17)). Hence Dworkin’s refutation does not apply.

In several respects the prima facie solution is a close relative of the positive/negative solution. According to both, where  $-S(p)$  &  $-S(-p)$  (or  $-S(Pp)$  &  $-S(Np)$ ) there is no legal gap, as the true/false solution dictates. Both therefore allow that certain legal propositions are true even without a source and both confine this to negative propositions ( $-Pf\ p$  or  $Np$ ). Yet the prima facie solution is not committed to a general categorization of all propositions into negative and positive. It relies instead on a distinction between internal and external negations of prima facie propositions and on the view that the absence of a prima facie legal state of affairs does not require a source.<sup>12</sup>

By introducing the notion of a prima facie proposition, however, this solution can dispense with theorem (12) which states that  $-(S(-p) \& S(p))$ . This theorem (and theorem (11)) cannot

<sup>12</sup> This absence of distinction between  $Pp$  and  $Np$  leads to some apparently counter-intuitive results such as (i)  $\vdash -S(Per\ x, \phi) \leftrightarrow -Pf(Per\ x, \phi)$ . That is, if there is no source permitting  $x$  to  $\phi$  then it is not the case that prima facie  $x$  is permitted to  $\phi$ . But this is not really paradoxical. Assume that (ii) either  $S(Pr\ x, \phi)$  or  $-S(Pr\ x, \phi)$ . If (iii)  $S(Pr\ x, \phi)$  then (iv)  $Pf(-Per\ x, \phi)$  and in such a case (i) is not paradoxical. If (vi)  $-S(Pr\ x, \phi)$  then (vii)  $-Pf(Pr\ x, \phi)$  (by (18)). According to (15) (when adapted to  $Pf$  propositions) (vii) entails  $Pf(Per\ x, \phi)$ . The prima facie permission results from the absence of a prohibition. That it does not result from the absence of an explicit permission is not on reflection surprising.

be derived from (16) and (18). All that follows from them is:

$$(19) \vdash (S(-p) \& S(p)) \rightarrow (Pf p \& Pf -p).$$

If the antecedent is true then there are conflicting legal rules, but no contradiction. The rejection of (11) and (12) seems to be an advantage. Whether or not the admission that the law may contain gaps is an advantage, it should be mentioned that the *prima facie* solution, while rejecting the claim that there are gaps where the law is silent, opens the possibility of a new kind of gap, i.e. through the existence of insoluble conflicts. If  $Pf p \& Pf -p$ , and if there is not sufficient ground to determine how the conflict is to be resolved, then neither  $p$  nor  $-p$  is the true conclusive legal position. In such a case we say that the law is unsettled and this means that it contains a gap.

## II. Sources and Laws

### A. Open Questions

The various solutions outlined in the above discussion give rise to many new questions. If the negative/positive solution is viable it must provide a test by which all legal statements can be identified as either positive or negative. If the true/false solution is to survive it must explain why it is thought that closure rules like: 'Whatever is not prohibited is permitted', 'If a contract is not legally valid then it is legally invalid', 'If a person is not guilty he is innocent', etc. are often thought to be universally valid. It is their supposed validity which provides the intuitive grounds for accepting the positive/negative solution. But the true/false solution is based on a rejection of these closure rules as universally valid (they may still be 'locally' valid, for some legal system) for it asserts that a man can be neither guilty nor innocent, a contract neither valid nor invalid, an act neither prohibited nor permitted. If the negative/positive and the true/false solutions are incompatible, much more needs to be said about their precise detail before judgment can be passed about the success of either of them.<sup>13</sup> The *prima facie* approach brings with it its own problems and in particular the problem of the relation between *prima facie* and conclusive judgments.

Beyond these, and similar questions which are specific to

<sup>13</sup> As will emerge below, the two solutions are compatible if properly interpreted.

different solutions, there are the more important questions which are common to all these solutions and to other possible ones: What is the exact nature of the relation between a law and its source? Does the sources thesis entail the possibility of gaps, and if so, on what grounds? How is the existence of gaps related to the existence or absence of jurisdiction to adjudicate in such cases?

Three solutions were outlined above to demonstrate the main routes a legal positivist may take and some of the main problems and consequences involved in each route. For the rest of the present essay a solution will be developed, different from those presented above, but essentially a variation of the *prima facie* solution. In developing it, it is hoped that some of the problems mentioned above will be at least partially answered.

## **B. Legal Statements**

Sources are normally thought of as sources of law. The occurrence of an event of a certain kind creates a law. Certain legal statements are true in virtue of that or other laws alone. These are pure legal statements. Other legal statements are true because of the existence of the laws and of other facts, such as the making of contracts, wills, the commission of offences and other wrongful acts, the exercise of public or private powers, the performance of duties. These are applied legal statements.

The usefulness of 'interposing' laws between sources and statements is self-evident. First, it distinguishes law-creating facts and other facts relevant for the truth of some legal statements, such as the commission of wrongs or of private transactions. It does, in other words, play a major role in forging the network of distinctions between private and public law. Laws are also one of the devices enabling us to refer to the content of the law without referring to the circumstances of its creation, the details of which are irrelevant for most practical purposes. When one wants to talk about what is commanded, rather than about who commanded and when and how, one refers to commands and all one needs to know is that they were properly issued in one way or another. For similar reasons we talk more often of laws than of legislative acts, etc., as imposing duties.

For present purposes, however, omitting laws and discussing

directly the relations between sources and legal statements simplifies matters without distorting any essential facts. But it is important to remember that 'sources' here means not just law-creating acts but all sorts of facts which make legal statements true or false. What, then, are legal statements? They are expressed in many forms, but most commonly through the use of the operators 'It is the law that...' or 'Legally...' or by the adjective 'legal' ('He has a legal right to...', etc.). Often no linguistic indication is given to show that a sentence is used to make a legal statement: 'You ought not to do that' may be used to make a legal statement. Ultimately it is impossible to identify legal statements by their form or through any linguistic indicators. They are identified by the nature of their truth-conditions. This point will be taken up again below. Though there is no linguistic form the use of which is necessary to express legal statements, all of them can be expressed by deontic sentences (i.e. sentences about what is or is not to be done, what rights, duties, permissions, liberties, powers people have or lack, what transactions were effected, etc.) preceded by 'Legally...' or 'It is the law that...' and sentences obtained from such sentences by the operations of sentential, quantificational, and modal logic. Such sentences can be viewed as the canonical form of legal statements. ( $Lp$  will indicate *Legally*  $p$  in formalized formulations.)<sup>14</sup>

A statement is not a statement of law in the abstract but a statement of English Law or of German Law, etc. So the operators should be 'It is English law that...', etc., but reference to the particular system can be omitted in the present discussions so long as it is remembered that all legal statements are statements of different particular legal systems.

Legal statements, like deontic statements generally, are logically stratified. Statements of legal reasons for action form the most elementary stratum. Permissions, duties, and powers are explained in terms of reasons; rights are explained in terms of duties, permissions, and powers; legal transactions in terms of rights and powers, etc. To analyse the nature of legal gaps in

<sup>14</sup> Note that 'there is a law that  $p$ ' entails 'legally  $p$ ' and normally implies that  $p$  is part of the content of a single law. As explained above, for the purposes of the present discussions such statements are assimilated to other legal statements and are not treated as a special category.

full, one has to examine each kind of statement separately. But in fact, once the problem of gaps is analysed for reason-statements, all the rest follows, given an adequate analysis of the other deontic concepts. The present discussion will therefore be confined to the case of reasons.

Three kinds of reason-statements will be discussed:

- (1)  $p$  is a reason for  $x$  to  $\phi$  or  $(p R x, \phi)$  for short.
- (2) There is a reason for  $x$  to  $\phi$  or  $(R x, \phi)$  for short.
- (3) There is a conclusive reason for  $x$  to  $\phi$  or  $(R_c x, \phi)$  for short.

Statements of kind (1) will not be analysed here. Statements of kind (2) are true if and only if a corresponding statement of type (1) is true. There could be a reason to  $\phi$  and at the same time a conflicting reason not to  $\phi$ . There is a conclusive reason for  $x$  to  $\phi$  if there is a reason for  $x$  to  $\phi$  which overrides all conflicting reasons, is not excluded by exclusionary reasons, nor cancelled by any cancelling conditions. There cannot be conflicting conclusive reasons; i.e.

$$(20) \vdash \neg(R_c x, \phi \ \& \ R_c x, \bar{\phi}).$$

At the same time it is not to be assumed that in every case to which reason applies there is a conclusive reason either for the action or against it; i.e.

$$(21) \text{ '}(R_c x, \phi) \vee (R_c x, \bar{\phi})\text{' is false.}$$

There are many kinds of permissions. For present purposes two kinds should be considered:

- (22)  $x$  has explicit permission to  $\phi$  ( $Per_e x, \phi$ ) = *df.* There is a fact with some force to cancel reasons for no- $\phi$ -ing ( $\bar{\phi}$ -ing).
- (23)  $x$  has a conclusive permission to  $\phi$  ( $Per_c x, \phi$ ) = *df.* It is false that there is a conclusive reason for  $x$  to  $\bar{\phi}$  (i.e.  $\neg R_c x, \bar{\phi}$ ).

A conclusive permission to act is the contradictory of a conclusive reason for refraining from that act. Hence the following is a logical truth:

$$(24) \vdash (R_c x, \phi) \leftrightarrow \neg(Per_c x, \bar{\phi}).$$

In the main the definition of a conclusive permission follows

the customary approach to permissions. The notion of an explicit permission is less familiar. Philosophers and logicians often think of 'permissions' as a nominalization of the deontic 'may'. But in ordinary discourse 'permission' means some act of granting permission and a person is permitted to act only if he is granted permission to do so. 'Explicit permission' is meant to explicate this notion of granted or obtained permissions. 'He allowed me to enter the room' imports that there was a reason why I should not which his permission cancelled. But a permission can be granted not only to cancel an existing reason but also to forestall possible reasons by cancelling them in advance: The manager and his secretary may object to a person's having access to the files. If either does object then the person has a reason not to approach the files. A journalist seeks to consult the files. He has no reason not to, since so long as no objection is made the files are available to the public. Anticipating that the secretary may object he obtains the permission of the manager. The permission cancels the force of the secretary's objection (if and when it materializes).

These are the only kinds of deontic statements to be discussed here. All other and more complex kinds of deontic statements are explained by pointing out their logical relations with these elementary forms of statement. Therefore though the following is a discussion of one particular kind of statement, the considerations invoked apply to all legal statements of all kinds.

### C. Sources as Reasons

Statements of the form  $LR x, \phi$ , i.e. there is a legal reason for  $x$  to  $\phi$  (which mean the same as 'Legally  $x$  ought to  $\phi$ ', 'It is the law that  $x$  ought to  $\phi$ '), are true, according to the sources thesis, because of the existence of a source, i.e. an appropriate social fact specifiable without resort to moral argument. But the relation between the statement of a source and the corresponding legal statement is not identity of meaning but that between ground and consequence. Two common answers to 'Why ought one to  $\phi$ ?' are 'Because there is a new law to that effect' or 'Because last year Parliament decreed so'. Both come to much the same thing (but see section IID.). Since reference to laws is not examined in this essay, only the second kind of answer is of present concern. It calls attention to the

very familiar fact that the making of contracts, marriages, laws, etc., are reasons for action. One ought to do as one contracted to do because one contracted to do so. One ought to maintain a certain person because one married him. One ought to pay income tax because Parliament enacted so, etc. Legal sources are reasons for action. If  $s$  is a variable ranging over statements of social sources, then  $sLRx, \phi$  (' $s$  is a legal reason for  $x$  to  $\phi$ ') is the general form of statements of legal reasons. Such statements are true if and only if  $s$  is true and is a statement of the appropriate social condition according to the doctrine of identity.

Because the existence of the appropriate source is a reason for  $x$  to  $\phi$ , its existence is the ground for the truth of statements of the form 'Legally  $x$  ought to  $\phi$ ' ( $LRx, \phi$ ). These are essentially existential statements asserting that there is some fact which is a legal reason for  $x$  to  $\phi$  ( $LRx, \phi$  is essentially the same as  $(\exists s) sLRx, \phi$ ). The existence of the source is the ground for the truth of a statement of what one ought, legally, to do, in the same way in which the existence of an appropriate instance is the ground for the truth of an existential statement. On this interpretation the sources thesis entails:

- (25) Statements of the form  $pLRx, \phi$  are true only if statements of social facts specifiable without recourse to moral argument are substituted for  $p$ .

The sources thesis entails a similar condition concerning explicit permissions:

- (26) Statements of the form  $LPer_x, \phi$  are true only if there is a social fact specifiable without recourse to moral argument which has the force to cancel legal reasons.

The facts alluded to are facts like an Act of Parliament repealing a legal prohibition or duty, a constitutional guarantee of certain freedoms, a licence to use a certain property given by one person to another. The sources thesis, it should be remembered, is a thesis about the law not about practical reason in general. Theses (25) and (26) are not meant as an assertion that only social facts can be reasons or permissions but that only such facts can be legal reasons and legal permissions.

It follows that while there is always a source for statements

of the form  $LRx, \phi$ , there is no source for the negative statements:  $\neg LRx, \phi$ ,  $L\neg Rx, \phi$ . Whether there is a difference between 'it is not the case that one legally ought to  $\phi$ ' and 'it is the law that it is not the case that one ought to  $\phi$ ' will be considered below. Intuitively, the negation of a legal reason is nothing more than the absence of a legal reason. It does not make sense to look for a source for the absence of a reason, legal or otherwise. This is but a reflection of the fact that  $\neg(p LRx, \phi)$  has no source either.

Here Professor Dworkin's refutation is avoided in a new way: by denying that there could be a source for a certain kind of legal statement. This claim may seem surprising and, therefore, a further explanation may be in place. It is often said, for example, that though a permission can be inferred from the absence of a prohibition, it can also result from a direct law granting it. Hence permissions can have sources. This, of course, has not been denied. What has been denied is that permissions of the same kind can be either based on a source or be sourceless. Explicit permissions are always source-based, conclusive permissions never are. The enactment of a law repealing a legal prohibition is an explicit permission:<sup>15</sup> it cancels a reason. The granting of a prospective permission confers an immunity: it is a fact with force to cancel the force of certain facts as reasons. A constitutionally guaranteed liberty gives one immunity against parliamentary legislation through having the force to cancel reasons arising out of parliamentary legislation. Normally when one has an explicit permission, one has a conclusive permission too. Hence often ordinary statements about freedoms, liberties, rights, and permissions do not, and have no need to, distinguish between these two kinds of permissions. Once the distinction is drawn it follows from the definitions that explicit permissions always depend on sources whereas conclusive ones never do.

Conclusive legal reasons are ordinary legal reasons, and thus are themselves social facts. They are undefeated reasons and thus depend also on the absence of defeating facts. The sources

<sup>15</sup> It may be worth noting that the cancelling permissions discussed in the first essay are a subclass of explicit permissions. Since, as was pointed out in the second essay, all legal reasons are protected reasons, all explicit legal permissions are cancelling permissions and the ability to grant them is a normative power.

thesis requires that the way conflicts of reasons are to be resolved is also determined, to the extent that it is legally determined at all, by social facts. Such facts are also legal sources.

#### **D. Ultimate Legal Rules**

The claim that legal sources are reasons for action raises as many questions as it solves. Are legal sources moral reasons or prudential reasons, or is there a special and distinct kind of reasons which legal reasons exemplify? Do ordinary legal statements import moral approval of the law? These questions cannot be explored here (they are indirectly answered in Essay 8 below). But suppose one asks why is a certain legislative act a reason for action? Is it not because of moral grounds that a policeman's order, for example, is a reason for action? Be that as it may, some of these grounds are legal while others may not be. The policeman's order is a valid reason because, generally, policemen act to preserve the peace and are reliable. This is not a legal ground. Another ground for accepting that the policeman's order is a reason for action is that Parliament conferred on him power to give such orders. There may or may not be non-legal grounds for accepting legal sources as reasons, but there are always such legal grounds. Or almost always, for in the end one gets to ultimate legal rules.

Suppose it is asked of English law: why is it true that parliamentary legislation is binding on the courts? The answer is that this is so because of the practice of the courts which follows a rule to that effect and because the rules practised by the courts of a legal system are rules of that system according to the doctrine of identity. Here (as in the case of all ultimate rules) the courts' practice is what makes the rule a legal rule and is thus its source. But the relation of source to rule and to the action the rule requires is different in the case of ultimate rules from that relation in the rest of the law.

With the rest of the law both the rule and its source could with equal justice be regarded as the reason for doing as the rule prescribes. Ultimate rules are likewise reasons for the action they require, but not so their source. That the English courts hold themselves bound to apply statutes is not the reason why they ought to do so. The rule that they should apply statutes is such a reason. The practice is no more than proof (constitutive proof) that the rule is a legal rule. It is neither a ground

for the validity of the rule nor for the action it prescribes. It is this fact which establishes the character of the rule as an ultimate *legal* rule. The fact that a rule is an ultimate legal rule means no more than that there is no legal ground, no legal justification for its validity. It does not imply that there is no ground or justification for the rule, only that if such ground exists it is not a legal one. With ordinary legal rules their source is the legal ground of their validity and a reason for behaving as they prescribe. That Parliament so enacted is a ground for the validity of the law and a reason for the required behaviour. These are legal reasons for their character as grounds of validity is itself determined by another law. By definition ultimate legal rules are not similarly grounded in legal reasons. The absence of a further law determining the grounds of validity of the ultimate rules is precisely what makes them *ultimate* legal rules.

Because further legal rules (themselves grounded in social facts) determine which facts create rules and are thus, with those facts, the grounds of validity of the rules, they can be used to identify the rules for the validity of which they are a ground. Since all the grounds are social facts or legal rules grounded in social facts, this identification is in accord with the sources thesis. There are no legal grounds for the validity of ultimate rules; no justification of them is provided by law. If they are identifiable by social facts, i.e. if they have sources, these facts cannot be the legal grounds for their validity. Therefore they cannot be legal reasons for doing as they require. They are identifying criteria only. The courts themselves use them as such. English courts look to their own practices when asking whether they are bound by a fraudulently obtained Private Act or by an Act of Parliament binding future Parliaments. But they refer to the source of the law in the practice of the courts in order to identify the precise content of the ultimate legal rule which binds them, not in order to justify it. Therefore, though the ultimate rule itself is the reason for the action it requires, its sources are not.<sup>16</sup>

<sup>16</sup> Hart was right in asserting that the rule of recognition, like all other rules, rests on social sources. Kelsen, however, was right in insisting that the relation between ultimate rules and their sources is different from that between ordinary rules and their sources. Positivists often claim that controversy is proof that the law is unsettled. By the analysis presented here, this is true of ultimate rules but not of others.

### III. Legal Gaps

#### A. The Nature of Gaps

There is a gap in the law when a legal question has no complete answer. Understanding a question is knowing what counts as a correct answer. This does not mean knowing which is the correct answer. It means knowing which statements are possible answers, i.e. which statements would be, if true, the correct answer. A legal question is a question all the possible answers to which are legal statements. A legal gap exists if none of the possible complete answers to a legal question is true.

Where a question lacks an answer it has various correct secondary answers such as ‘This question has no answer’, ‘There is no law on this matter’, ‘The law is unsettled’, etc. It also may have partial answers. This is so in particular where some of the possible complete answers are false while the others are neither true nor false. The existence of partial answers and of secondary answers is compatible with the existence of a legal gap, i.e. with the absence of a complete answer.

A court may or may not have jurisdiction over a given legal question. It has jurisdiction over a question when it is entitled to look for its answer and rely on it in rendering decision. Essentially this means that it has jurisdiction when its decision as to the right answer to the question is *res judicata* (if the answer is an applied legal statement) or when it is the basis on which the courts proceed to create the *res judicata* (if the answer is a pure legal statement). On the basis of this definition we can distinguish two kinds of gaps in the law:

*Jurisdictional Gaps.* A legal system is jurisdictionally complete if its courts have jurisdiction over all legal questions. It has a jurisdictional gap if its courts lack jurisdiction over certain legal questions.

*Legal Gaps.* A legal system is legally complete if there is a complete answer to all the legal questions over which the courts have jurisdiction. It contains a legal gap if some legal questions subject to jurisdiction have no complete answer.

It is possible, of course, to have a wider concept of legal gaps, according to which the mark of a gap is the existence of any

unanswered legal question. But the narrower concept defined here is the more important one. Gaps present legal and philosophical problems only if they arise from questions over which there is jurisdiction. A legal system is a normative system providing for its own application through judicial bodies, and judicial bodies play a crucial role in any doctrine of identity for legal systems. Since the law is identified through the eyes of the courts, legal gaps should be so identified too. Yet 'a legal question' should not be identified as any question a court may answer in rendering decisions. It is not to be taken for granted that all the questions courts are concerned with are legal questions. A legal question is identified independently by the character of its answers: only legal propositions are possible complete answers to a legal question and legal propositions of a given legal system are identified by their truth-conditions as determined by the doctrine of identity.

In every case there is one legal question which encompasses all the others: What decision does the law require in this case? A complete answer is provided by pointing to a decision which the law requires. When no decision is required by law the question is unanswerable and there is a legal gap. In such a case the question will have secondary, and possibly partial, answers only, such as 'the law is unsettled', 'the law requires nothing'. It is here assumed that the question is not addressed to the court's own jurisdiction and therefore 'the law requires a decision that there is no jurisdiction' is not an answer to this question.

When the court's decision turns on whether or not a person ought (conclusively) to perform a certain action, there are two possible complete answers:

- (1)  $LR_c x, \phi$ , the law conclusively requires that action; and
- (2)  $LPer_c x, \bar{\phi}$ , i.e. the law conclusively permits the omission of that action.

If neither proposition is true there is a legal gap for there is no decision required by law. It follows that there are two possible kinds of legal gaps:

- (1) ' $LR_c x, \phi$ ' is neither true nor false and ' $LPer_c x, \bar{\phi}$ ' i.e. ' $L - R_c x, \phi$ ' is neither true nor false.
- (2) ' $(-LR_c x, \phi) \& (-LPer_c x, \bar{\phi})$ ' is true.

Both are situations of gaps if the questions to which the relevant propositions are answers are (at least if they have complete answers) within the jurisdiction of the courts.

## **B. Indeterminacy and Gaps**

Whether or not there is a legal reason to perform a certain action is a matter of identifying which legal reasons there are and whether they are reasons for that action. Ultimately it is a question of interpreting the precise import of certain facts such as acts of legislation, judicial decisions, etc. Whether a certain executive act is a legal reason (i.e. whether it creates a binding legal norm) is a matter of interpreting a parliamentary legislative act which conferred legislative powers on the executive as well as of interpreting the executive act itself. The same is true of deciding whether that executive act is a reason for a certain action. To a considerable extent these questions revolve round the interpretation of the language used and the intentions with which certain acts were performed. What do the words of the contract, will, company memorandum, or statute mean and what did those who formulated them intend, are the most familiar (though not the only) questions involved in identifying and interpreting legal reasons.

Questions of intention and meaning may have no answer. It may be neither true nor false that in making the remark a person intended to insult his colleague, and then it is also neither true nor false that he did not so intend. Similarly it may be neither true nor false that this is a dwelling or a motor vehicle. It may be neither true nor false that this man is bald. Where the facts which are legal reasons are indeterminate, through vagueness, open texture, or some other factors, certain legal statements are neither true nor false.<sup>17</sup> Such indeterminacy may make it neither true nor false that this

<sup>17</sup> Normally indeterminacy of intention creates a legal gap only if reflected in the use of indeterminate language. Closure rules—on which see below—come into operation in most other cases. Sometimes, however, gaps arise because of indeterminate intentions independently of the language employed. Suppose that a statute requires ‘ship-owners’ to  $\phi$ . Years after its enactment a new kind of vessel is developed which is beyond doubt a ship, but it is neither true nor false that the legislator intended the obligation to apply to the owners of such vessels. Many legal systems hold that the plain language of the statute is at least a prima facie reason for such owners to  $\phi$ . But a system’s rules of interpretation may make such questions depend exclusively on the legislator’s intention (and they may regard the language employed merely as presumptive evidence of its intentions). In such a system the law on the question is indeterminate.

structure is liable to tax as a dwelling, etc. A cause of legal gaps of the first type is the indeterminacy of language, of intention, and of other facts. It is worth noting that this kind of legal gap is not the law's peculiarity. They are totally dependent on and derive from gaps in statements of an ordinary and not a particularly legal kind such as statements about intentions and language. It is the indeterminacy of ordinary everyday facts which generates legal gaps.

Are such gaps inevitable? It seems that the sources thesis makes them unavoidable since it makes law dependent on human action with its attendant indeterminacies. In 'No Right Answer?' Professor Dworkin denies that vagueness inevitably leads to gaps. In part, his view presupposes a rejection of the sources thesis and does not therefore affect the present argument. But, in part, he relies on the possibility of rules of interpretation for handling vague terms such as 'if a statute uses vague language it must be taken to have changed the legal *status quo ante* only to the extent justified by the indisputable core of the language employed' (p. 68). This suggestion rests, however, on a fallacious view of vagueness. It assumes that whereas a term which is not vague divides all cases into those to which it applies and those to which it does not, a vague term divides all cases into three sets: those to which it applies 'by its indisputable core meaning', those to which it clearly does not apply, and those in between. It is as if a term is vague because it draws two sharp dividing lines instead of one. The truth is that all, and not only some, nouns, verbs, adverbs, and adjectives of a natural language are vague. And though a vague term clearly applies to some cases, clearly fails to apply to some, and doubtfully applies to others, yet it is often impossible to draw general boundary lines between the three categories. It is a test of adequacy of any account of vagueness that it recognizes as a central type of it the cases where vagueness is 'continuous'.

Terms are vague in this sense if besides cases on the borderline between the area covered by the concept and that which it does not cover there are borderline cases between those covered by the term and those which belong to the borderline and there are borderline cases between those to which the term does not apply and those on the borderline, and so on indefinitely. With respect to a vague term of this kind one is on occasion uncertain



conflicting reasons is conclusive. One is where the two are equally balanced. They cancel each other and it is false that there is a conclusive reason for the act and false that there is a conclusive reason for its omission.  $(L - R_c x, \phi) \& (L - R_c x, \bar{\phi})$  is true and it follows that  $(LPer_c x, \bar{\phi}) \& (LPer_c x, \phi)$  is true. This kind of situation involves no unresolved conflict nor any legal gap.

Completely different is a situation of unresolved conflict. It arises when conflicting reasons fail to override each other, not because they are equally matched, but because they are not matched at all: for whatever reason, the conflicting reasons are incommensurate as to strength. Neither is stronger nor are they equal in strength. In such a case it would be wrong to say that the agent is permitted to perform the act. But it would be equally wrong to say that he is not permitted to perform it. The only intuitively correct description of his state is that he is neither permitted nor not permitted to perform the act. This is possible only if the statement that he is permitted as well as its negation are neither true nor false. It follows that  $LR_c x, \phi$  and  $LR_c x, \bar{\phi}$  are equally neither true nor false. Hence unresolved conflicts give rise to legal gaps and these, like all other gaps due to indeterminacy in the sources, are gaps of the first kind.

There is yet a third way in which the sources thesis is responsible for legal gaps and it too arises out of conflict situations. The law may make certain legal rules have prima facie force only by subjecting them to moral or other non-source-based considerations. Let us assume, for example, that by law contracts are valid only if not immoral. Any particular contract can be judged to be prima facie valid if it conforms to the 'value-neutral' conditions for the validity of contract laid down by law. The proposition 'It is legally conclusive that this contract is valid' is neither true nor false until a court authoritatively determines its validity. This is a consequence of the fact that by the sources thesis courts have discretion when required to apply moral considerations.

#### D. Closure Rules

Because of the sources thesis indeterminacy in sources always leads and absence of sources for sufficient conflict reserving rules may lead to legal gaps of the first kind. Can there be legal gaps of the second kind? Gaps of the second kind

can exist only if it is possible that  $(-LR_c x, \phi) \ \& \ (-LPer_c x, \bar{\phi})$ . That is, if it is possible that

$$(27) \quad (-LR_c x, \phi) \ \& \ (-L-R_c x, \phi).$$

But it is a logical truth that

$$(28) \quad \vdash \quad -(LR_c x, \phi) \leftrightarrow (L-R_c x, \phi).$$

From (27) and (28) follows

$$(29) \quad (L-R_c x, \phi) \ \& \ (-L-R_c x, \phi).$$

(29) is a contradiction. Hence (28) entails the rejection of (27) and with it of the possibility of gaps of the second kind. The derivation of (28) is simple. First,  $L-R_c x, \phi \rightarrow -LR_c x, \phi$ . It cannot be true that legally there is a conclusive reason to perform an act and at the same time that legally there is no such conclusive reason. Therefore

$$(30) \quad \vdash \quad LR_c x, \phi \rightarrow -L-R_c x, \phi.$$

$L-R_c x, \phi \rightarrow -LR_c x, \phi$  follows from (30) by counterposition. Next the derivation of  $-LR_c x, \phi \rightarrow L-R_c x, \phi$ . Assume that the antecedent is true, i.e. that  $LR_c x, \phi$  is false. Can it also be that  $L-R_c x, \phi$  is false? If  $-LR_c x, \phi$ , then it is not the case that there is a fact which is a conclusive legal reason to  $\phi$ . But if there is no conclusive legal reason to  $\phi$ , can it fail to be the law that there is no conclusive reason to  $\phi$ ? This would have been a real possibility had there been a need for a special kind of fact to make true negative statements of reasons of the form  $-R x, \phi$  (as was confusedly imagined to be possible by the negative/positive solution. Cf. (7) above). But it has already been established (in section IIB) that such propositions are verified by the absence of reasons only and it is precisely that absence which is entailed by  $-LR_c x, \phi$ . Hence  $-LR_c x, \phi \rightarrow L-R_c x, \phi$ , Q.E.D.

(28) is logically equivalent to (31):

$$(31) \quad \vdash \quad (-LR_c x, \phi) \leftrightarrow (LPer_c x, \bar{\phi}).$$

This is a rendering of the familiar closure rule that whatever is not legally prohibited is legally permitted and vice versa. So far as (31) is concerned both implications are equally valid. The primacy of the implication from left to right (i.e. whatever is

not prohibited is permitted) is due to the positive/negative distinction underlying the whole position, i.e. one needs some facts to have reasons, but none is needed for there not to be a reason. Thus the present approach validates this and several other closure rules and by accepting them rejects the possibility of the second type of legal gaps.

The outcome of this discussion is that legal gaps are not only possible but, according to the sources thesis, inescapable. They arise, however, where the law speaks with an uncertain voice (simple indeterminacy) or where it speaks with many voices (unresolved conflicts). Contrary to much popular imagining, there are no gaps when the law is silent. In such cases closure rules, which are analytic truths rather than positive legal rules,<sup>19</sup> come into operation and prevent the occurrence of gaps.<sup>20</sup>

<sup>19</sup> They are analytic truths concerning *legal* propositions derivable from the sources thesis. They cannot be derived in normative systems where such a thesis does not hold.

<sup>20</sup> I am grateful to Simon Blackburn, Ronald Dworkin, and Mary Tiles for instructive comments on earlier drafts of this essay.

## The Identity of Legal Systems\*

Laws are part of legal systems; a particular law is a law only if it is part of American Law or French Law or some other legal system. Legal philosophers have persistently attempted to explain why we think of laws as forming legal systems, to evaluate the merits of this way of thinking about the law, and to make it more precise by explicating the features that account for the unity of legal systems. Various theories have been suggested but none has been accepted as completely satisfactory, and the continuing debate owes much to the intricacy of the problems involved. This intricacy is partly due to a considerable cloudiness in the understanding of the problems themselves. I will attempt, through this essay, to clarify the nature of the problem of the unity of municipal legal systems. My primary aim is not to answer all the relevant questions, but rather to formulate them more precisely, for I believe that in philosophy a clear conception of a problem is half the way to its solution.

The term 'a legal system' is not a technical legal term. It may occasionally figure in courts' decisions, but any term can appear there. Although it is occasionally used in legal argument, it has not the character of a technical legal term such as 'floating charge' or 'fee simple' or 'consideration'. Nor is the concept important to the day-to-day administration of law, as are the concepts of contract, ownership, right, duty, and the like. The term is primarily used in thinking about the law, not in the actual use and application of the law. It is commonly used in books of jurisprudence or comparative law, not in books about property law, torts, or copyright.

Therefore, when trying to clarify the notion of a legal system, the legal theorist does not aim at defining clearly the sense in which the term is employed by legislators, judges, or lawyers.

\* First published in *California Law Review* (1971). I am indebted to Professor H. L. A. Hart for many discussions on the problems dealt with in this essay and for his illuminating comments on a first version of it.

He is, rather, attempting to forge a useful conceptual tool, one which will help him to a better understanding of the nature of law. This does not mean that he should not try to model the concept in a way that would be useful to the solution of certain legal problems. Rather, it means only that even if he does this, he will not be trying to elucidate the meaning of a technical legal term, but instead to provide the legal practitioner with a concept that may help him in tackling some nagging legal problems. But should it really be the goal of the legal theorist to help solve legal problems by moulding a precise concept of the identity of a legal system? To what kind of legal problems may such a concept be relevant? And if this is not his goal, what is? Providing an answer to these questions is the main purpose of this essay.

To try to answer these questions we should first distinguish between two possible concepts of the unity of a legal system, which I call the material unity and the formal unity. The material unity of a legal system consists in its distinctive characteristics; it depends on the content of its laws and on the manner in which they are applied. When trying to explain the characteristic features of a legal system we are not, of course, looking for the detailed regulation of every legal institution. Rather, we are looking for the all-pervasive principles and the traditional institutional structure and practices that permeate the system and lend it its distinctive character. To distinguish the problem of formal unity from that of material unity, I shall call the former the problem of the identity of legal systems.

## I. The Problem Isolated

The identity of the system is found in the criterion or set of criteria that determines which laws are part of the system and which are not. Some proposed answers to the problem of identity are well known; perhaps the best-known are: A law is part of a legal system if and only if it was enacted directly or indirectly by the sovereign of that system (Austin), or if and only if it is authorized by the basic norm of the system (Kelsen), or if and only if it ought to be recognized according to the rule of recognition of the system (Hart). These three philosophers were not concerned with the material unity of legal systems. They did not think that the unity of the system depends on the

content or spirit of its laws, or on the traditions and practices of its most important legal institutions. Instead, they hoped to formulate a test that would enable them to determine whether any two laws belong to the same legal system or not.

A more or less clear concept of the identity of a legal system is presupposed by any investigation into its material unity. The investigator needs to know which laws and institutions form the system before he can inquire into its distinctive characteristics. He does not need, however, a complete and precise list of the laws of the system. It suffices for his purposes to be able to identify the bulk of the laws. Indeed, the investigation into the material unity of a legal system may even help to decide some borderline cases concerning its identity. When all other indicators fail to provide an answer, it may be reasonable to decide that a norm is part of a certain legal system on the ground that it closely resembles in spirit and manner of application the rest of the laws in that system, or that since it is diametrically opposed to the character of the system it should not be regarded as a part of it.

In describing the problem of identity as a quest for criteria that determine which laws belong to a legal system, it is normally assumed that the notion of a law is clear and uncontroversial. In fact this is far from being the case. The problem of the individuation of laws—the question of what is one complete law—is one of the most controversial in jurisprudence. It would, therefore, be desirable to separate the problem of identity of legal systems from that of the individuation of laws. This can easily be achieved by reformulating the problem in the following way: The problem of identity of legal systems is the quest for a criterion or set of criteria that provides a method for determining whether any set of normative statements is, if true, a complete description of a legal system.

With the help of certain technical terminology, it is possible to side-step the problem of individuation. A statement is a normative statement if and only if the existence of a norm is a necessary condition for its truth. A normative statement is pure if and only if the existence of certain norms is sufficient for its truth. The set of all the pure statements referring to one legal system is called the 'total set' of that system, and every set of pure statements that is logically equivalent to the total

set of a system is a complete description of that system. The problem of identity is a search for criteria for a complete description of any legal system, and it is irrelevant whether any statement in the description describes just one complete law or not. Therefore, a theory concerning the nature and principles of individuation of laws need not be presupposed in an examination of the problem of identity.

By stipulating that the criteria identify sets of statements that are complete descriptions *if true*, the problem of the existence of legal systems may be separated from the problem of identity. It is the business of criteria of existence of legal systems to provide a method for determining when a complete description is a true description—to determine whether the legal system described actually exists. This task should be clearly separated from that of formulating the criteria of identity.<sup>1</sup>

The problem of identity has two quite distinct aspects: the aspect of the scope of a legal system and the aspect of its continuity. Questions of scope arise when we consider whether the conventions of the constitution, a valid contract, the regulations of a limited company or of a trade union, for example, are part of the legal system. Questions of continuity concern the various ways in which a legal system ceases to exist and is replaced by a new system. Does, for example, a revolution, or a *coup d'état*, or a declaration of independence, terminate the existence of one legal system and signal the emergence of a new system?

A momentary legal system is a legal system at a particular point of time. The problem of scope is the search for criteria of identity of momentary legal systems, whereas the problem of continuity is the search for criteria providing a method for determining whether two momentary legal systems are part of one, continuous, legal system. I will show that somewhat different considerations apply to each of these problems.

These preliminary clarifications make it clear that problems of identity are here considered as jurisprudential problems, the solution of which lies in providing sets of criteria of identity, which, if successful, provide methods of determining the identity of *all* municipal legal systems. Thus conceived, the problem is very different from that facing a legal practitioner looking for an answer in a *particular* legal system to a certain

<sup>1</sup> Cf. J. Raz, *The Concept of a Legal System* (Oxford, 1970), pp. 49, 70.

legal problem. Consider, for example, a judge who must render a decision in a particular case before him. Some answers that might satisfy the legal philosopher will not help the judge. A legal philosopher may say, and some philosophers have said, that what judges do about disputes is the law; but this is unlikely to be of much help to a judge wondering what he should do about a dispute. Does this prove that this jurisprudential criterion is false or pointless? The answer to this question depends on one's position on the fundamental problem posed above: What purpose do legal philosophers have in constructing theories concerning the identity of legal systems? They do not mean merely to elucidate the current meaning of a legal term; what, then, is their purpose?

One possible answer is that the goal is to forge a concept that will help the courts and other people concerned with the law to find answers to certain legal problems. There is a sense in which this is trivially true of all jurisprudential problems. Jurisprudence strives to improve our understanding of the law, and in one way or another, however remotely or indirectly, an improved understanding of the law is bound to affect the operation of the law and to help legal practitioners. But the above-mentioned contention should not be taken in this sense, for its intent is to claim that there are certain specific legal problems the solution of which will directly benefit from a jurisprudential analysis of the identity of legal systems.

Several legal problems can be thought to be relevant to the concept of identity. Some relate to the scope of the legal system, that is, to the problem of the identity of momentary systems. Most legal systems distinguish between procedures of proving matters of fact and procedures of arguing about points of law before the courts. Many legal systems stipulate that ignorance of law is no excuse for committing an offence whereas (reasonable) mistake of fact is. Many legal systems include laws to the effect that a general law is valid only if published in a manner specified by law. All of these legal provisions can give rise to disputes turning on the question: Is some rule part of the legal system or not? Should the defence have proved the rule as a fact or is the rule part of the legal system? There are other doctrines that may also bring the question of the scope of the legal system before the courts. For example, are internationally

binding treaties to which the state is a party, part of the municipal law of that state? Are the rules of public international law part of the law of the land?

Questions pertaining to the continuity of legal systems may arise before the courts in different contexts. A successful *coup d'état* or the establishment of a new state may give rise to various problems: Are previous laws still in force? Do persons who previously held high office still hold it or should they be re-nominated? Can the new regime or state claim taxes and debts owed to the old one? Can a person who committed an offence before the change be prosecuted after it has occurred? And so on.

Some, perhaps most, of the problems mentioned have obvious solutions and present no difficulty to the lawyer or judge. It is the nature of the question, the considerations relevant to a correct answer, that matter. Its degree of difficulty is immaterial. Nor is it part of the claim that such problems are always decided according to the courts' conception of the scope or continuity of a legal system. Other legal considerations may be involved, and the courts may explicitly or implicitly decide the issues on the other grounds.<sup>2</sup> It follows that even when faced with such problems, the courts' opinions on the question of identity are not always clear from the actual decision in the case. The arguments used to justify the decision are a much better indication and even they may be open to different interpretations.

It may be suggested that legal theorists, when dealing with the problem of identity, should aim to help formulate a systematic answer to legal problems of the kind mentioned. They should take account of existing legal solutions. They should, for example, regard the fact that foreign law, when applied by courts according to the rules of private international law, must be proved by expert witnesses as an indication that it is not part of the legal system under which these courts operate. Legal theorists should generalize such particular legal solutions, elucidate the reasons for their adoption, and on this basis formulate a theory of identity that will help solve difficult and novel

<sup>2</sup> For example, the courts may decide that a certain law is part of the system and is valid, and the provision making validity dependent on publication should be construed to include an exception for laws of this kind.

problems of this kind as well as provide guides for evaluating the merits of accepted solutions.

In pursuing such a task the theorist may find himself torn between two conflicting considerations. On the one hand, he is striving to reveal an underlying unity in the solutions, whether accepted or proposed, of many diverse legal problems. On the other hand, he must face the fact that each solution bears different legal consequences and arises in a different context, and therefore each solution might well be guided by different considerations. The purpose of making validity depend on publication may be different from the law's purpose of declaring that a mistake of law is no excuse. The theorist cannot take for granted that examination of the various legal problems relating to the identity of the legal system will lead to any unified criterion of identity. Still less can he assume that conclusions reached in the study of one legal system will be applicable to any other legal system. Different legal systems uphold different ideals of justice, maintain a different balance between conflicting interests, and pursue somewhat different goals. They exist in different societies living under different conditions. They are, therefore, likely to adopt different solutions to the problem of identity.

These remarks do not show that an investigation into the various legal criteria of identity accepted in a certain legal system and their underlying rationale is not worth pursuing. Rather, they merely warn against hasty generalizations and against uncritical application of accepted solutions to novel cases. They also draw attention to the fact that while pursuing such an investigation the theorist is engaged to a considerable extent in a critical task. He does not merely enumerate legal criteria, he rather inquires into their justifications and on this basis attempts to reach a sound generalization that will help solve new or difficult cases reasonably.

It is probably these features of this particular purpose of a doctrine of identity that account for the fact that none of the great positivist legal philosophers interested in problems of identity regarded their inquiries as designed to further this purpose. Positivist jurisprudence conceived the problem of identity as a descriptive problem of general jurisprudence. They wanted to find criteria for identifying the laws of any legal system; they

were not interested in the correct or desirable solution to any set of specific legal problems. Bentham, Austin, Gray, The American Realists, Kelsen, Hart—none of them had any great interest in any of the legal problems enumerated above. They did not regard them as relevant to their theories because they were pursuing a different goal. The attempt to achieve a reasonable and systematic solution to a set of related legal problems of the type enumerated is a possible aim of a legal theorist. It is not, however, an aim that has been important in the history of jurisprudence.

## II. The Problem Clarified

What, then, were the aims of legal philosophers when dealing with problems of identity? There are three main issues that bear on the problem when considered as one of analytic jurisprudence:

- First, the relation between the existence of a law and its efficacy;
- second, the distinction between making a new law and applying an existing one; and
- third, the relation of law and the state.

Together these issues determine the answer to the problem of identity. Although each one of them has engaged the mind of some of the legal philosophers who have discussed the problem of identity, failure to consider all three partially accounts for the fact that no satisfactory solution to the problem has been found. This section is an examination of the three issues.

### A. The Relation of Existence and Efficacy of Laws

The question of the relation of existence and efficacy of laws is one of the most fundamental questions concerning the nature of law. It concerns the conditions for the existence of laws, but since laws exist only in legal systems, to ask whether a law exists is to ask whether it is part of the legal system concerned: the question refers to the problem of identity. There are two extreme positions on this issue. At the one extreme is the claim that a law created in the appropriate manner exists and is valid; its efficacy or inefficacy does not affect its existence and validity

unless another law of that system makes efficacy a condition of continued existence of laws. Diametrically opposed to this is the argument that laws exist because and to the extent that they are socially accepted and followed; social customs are laws even if not enacted, whereas enacted law is not valid if it has no roots in social practices. Some theorists offer various compromise solutions. Kelsen, for one, says:

A general legal norm is regarded as valid only if the human behaviour that is regulated by it actually conforms with it, at least to some degree. A norm that is not obeyed by anybody anywhere, in other words a norm that is not effective at least to some degree, is not regarded as a valid legal norm. A minimum of effectiveness is a condition of validity.<sup>3</sup>

Lasswell and Kaplan also prefer a compromise solution, although placing a much heavier stress on efficacy than does Kelsen's solution. Lasswell and Kaplan argue: 'Laws are not made by legislatures alone, but by the law-abiding as well: a statute ceases to embody a law (except in a formal sense...) in the degree that it is widely disregarded.'<sup>4</sup>

There is an element of truth in both views. Laws guide human behaviour, help people in planning and deciding on their future course of action, and provide standards for evaluating past or planned actions. A law, the existence of which is unknown, or that is never acted on by the police nor enforced by judges or juries, does not guide the behaviour of most people, not even that of law-abiding people. There seems, therefore, to be no reason to regard it as part of the legal system, since its complete inefficacy has deprived it of the main characteristic of law, that of guiding behaviour. Furthermore, not only do parliaments often modify the law in response to a change in social practice and to conform with prevailing opinions and customs, but some social practices that were not adopted by the legislature have characteristics of laws—they guide behaviour and very often they affect courts' decisions, sometimes without the judges even being aware of this.

As against these arguments, proponents of the opposite view argue that although both laws and social practices guide behaviour,

<sup>3</sup> H. Kelsen, *The Pure Theory of Law* (2nd ed., Berkeley, 1967), p. 11.

<sup>4</sup> H. Lasswell and A. Kaplan, *Power and Society* (New Haven, 1950), p. 75.

this is no reason to amalgamate the two. The concept 'a law' was traditionally used to refer to norms, whether or not they conform with social practices, that have the characteristic of being part of a system of norms united by its relation to legal institutions. Social practices differ from laws because they are not institutionalized. The contention that laws are laws because they were enacted by legal institutions clarifies this basic fact about the law, focusing attention on its institutionalized nature. Even legal custom is not law until it is recognized and declared to be law by the courts. Efficacy, therefore, does not affect the validity and existence of laws. To claim otherwise is to confuse law with social customs and to disregard the basic fact about the law—that it is created by institutions. There is, no doubt, a relation between law and social practices, but it is at most a causal relation. Legislatures and courts may, consciously or unconsciously, be affected in making new laws by prevailing customs and practices; the existence of laws may also affect social customs and habits and even modify opinions. There is, however, no logical connection between law and social practice.

Both arguments are attractive, and it is very tempting to accept a compromise solution, as Kelsen, Lasswell, Kaplan, and others have done. There is, however, yet another possible solution to the problem, which is not a combination of the first two yet incorporates the sound elements of both. This solution shifts the emphasis on to the law-applying institutions, and makes recognition by law-applying organs a necessary condition of the existence of laws. This in turn makes the institutionalized nature of law an indispensable part of the criteria of identity: a law is part of the system only if it is recognized by legal institutions. The emphasis is, however, on the law-applying rather than the law-creating institutions.

Such an approach may be justified by three reasons, the third of which underlies the other two. First, although law-creating institutions are of the greatest importance in modern societies, where law is conceived as the outcome of deliberate human decision as to what society should be like, they played a minor role or did not even exist in primitive societies, where the laws were conceived as immutable and were in fact changed mainly by slowly evolving customs. Law-applying institutions are, on

the other hand, a constant feature of law in every type of society and their existence should be regarded as a defining characteristic of law.

Secondly, since most legal systems recognize diverse sources of law, the only way to determine which are the law-making institutions and procedures of a given legal system is to establish which sources of law are recognized by the courts. Hence only the courts of a legal system can provide the clue to its criteria of identity.

Thirdly, it is an essential feature of legal systems that they are institutional, normative systems. It is, therefore, reasonable to take the law to consist of those norms, rules, and principles, that are presented to individuals and institutions as guides to their behaviour by the body of legal institutions as a whole. When the actions of law-creating and law-applying organs conflict, the actions of the law-applying organs are those that affect the considerations of the law's subjects; they have final authority to declare what is the law.<sup>5</sup>

The third rationale also preserves the importance of efficacy to the definition of law. Efficacy, however, is relevant only in so far as it affects the practices of the law-applying institutions. If, for example, the courts consistently refuse to act on a law, that law is not part of the legal system the courts operate, despite the fact that it was lawfully enacted and was never repealed. If the courts consistently interpret a statute in a way deviating from its original meaning, their reading of it, not its original sense, becomes the law. According to this approach, then, the existence of the law is logically related to the practice of the law-applying organs. The condition of a law's membership in a legal system is, however, a counterfactual: if presented with the appropriate case the courts would act on the law. This may be true even though they are never—or seldom—presented with the appropriate case. Prosecutions in criminal cases may seldom be made, and civil cases may always be settled out of court in a way contrary to the law. Therefore a law may be valid even though it is largely inefficacious.

These remarks do not spell out any solution to the problem of the relation between efficacy and existence of laws. Rather, they merely delineate an approach to the problem, defining a

<sup>5</sup> See generally J. Raz (n. 1 above) pp. 191–2, 201–2.

certain type of possible solution.<sup>6</sup> A large number of authors—among them Holland, Gray, Salmond, Holmes, Llewellyn, and Hart—opted for this approach and adopted some solution that comes within its scope, although they differed greatly in the details of their theories as well as in the reasons that led to their adoption. The outline of an argument for some solution such as is offered above does not represent the reasons of all of them; it may not even represent the reasons of anybody but myself. But what we all have in common is the emphasis on law-applying organs in our criteria of identity of legal systems.

Since the aim of this essay is to clarify problems rather than examine or propose detailed solutions, there is no need here to compare the various solutions to the problem of identity that fall within the delineated approach. Attention should, however, be drawn to two claims that, although made by some of the above-mentioned writers, were rejected by others, and are not a necessary concomitant of all theories of identity of this type. First, to claim that a law is part of a system only if it is acted on by the law-applying organs does not entail that these organs create the law. They may be, and on most occasions are, merely recognizing and enforcing laws previously created by legislation, precedent, or custom. Secondly, that a law is part of a system only if recognized by the courts does not entail that laws are descriptions or predictions of what the courts are doing or will do. Courts are composed of human beings and the causes of their actions are open to analysis and prediction by psychologists, sociologists, and other scientists just as much as are the causes of other people's actions. Their judgments are, however, the fruits of deliberate decisions based on the evaluation of reasons for the various alternatives. Moreover, the courts write down the reasons that, in their opinion, justify their decisions, and it is by examining the courts' opinions that one finds the laws on which they act. The laws themselves are, therefore, normative. They guide the actions of the courts as much as those of ordinary people. Their existence is ultimately based on

<sup>6</sup> Various problems are not even mentioned here. What is a law-applying organ? What is a court? Are the actions of all law-applying organs relevant to the criteria of identity? I refer in the text sometimes to courts, sometimes to law-applying organs, but I do not wish to express any opinion on these questions here. See further Essay 6 below.

social practices, but this is common to all positive norms and does not detract from their normativity.

## **B. The Distinction Between Making a New Law and Applying an Existing One**

The statement that a law is part of a legal system only if it is recognized by the law-applying organs—the courts—of the system means only that it would have been acted on by the courts had they been presented with the appropriate problem. That a court would apply a law if faced with a case to which the law applies is an indication that either the law exists in the legal system or that the law will be made by the courts when they have an opportunity to do so. Recognition by the courts or other law-applying organs is not a complete criterion of identity because these organs often have power to make new laws, and often what law they are going to make can be determined in advance. As a first step towards completing the criterion,<sup>7</sup> one must incorporate in it reference to the fact that the law would not only be recognized by the courts but would be recognized as a previously existing law. It is not a new law that they would make when faced with an appropriate case. For this reason the distinction between applying an existing law and creating and applying a new one is the second jurisprudential issue involved in the problem of identity. Of the major legal theorists Hart is the only one to face this problem, and a brief discussion of the relevant aspects of his theory will clarify the nature of the issue.

Hart argues that the distinction between the application of a new law and the application of a previously existing one turns on the existence or absence of a duty to apply the law. If and only if the court applies a law that it is under a duty to apply is it acting on a previously existing law; on the other hand, when it applies a rule that it has no duty to apply it is not acting on a previously existing law.<sup>8</sup> This is a consequence of Hart's

<sup>7</sup> Hypothetical recognition by the courts is not, strictly speaking, even a necessary condition for a law's membership in a legal system, for a law may belong to a legal system even if the courts have power to change it and it is known that they will do so, given an appropriate opportunity. The precise relation between the courts' practices and the identity of the legal system is clarified by the discussion in this part of Hart's theory of the rule of recognition. See pp. 91–5 below.

<sup>8</sup> In such cases the court transforms the rule into a law of the system if and only

doctrine of the rule of recognition. In every legal system, he argues, there is of necessity a rule of recognition that identifies the laws of the system;<sup>9</sup> the criterion of identity of legal systems can be formulated as follows: A legal system consists of a rule of recognition and all the laws identified by that rule. Hart's discussion of the rule of recognition falls short of the high standard of lucidity characterizing the rest of his book and requires interpretation, which will be limited to the doctrine's effects on the problem of identity.<sup>10</sup>

A rule of recognition is 'a rule for conclusive identification of primary rules of behaviour'.<sup>11</sup> Here, as occasionally elsewhere, Hart states that the rule identifies only primary rules. It is quite clear, however, that his rule of recognition is a rule for the identification of all the other rules of the system, and only them. It specifies 'some feature or features possession of which by a suggested rule is taken as a conclusive affirmative indication that it is a *rule of the group*'.<sup>12</sup> This means that the rule of recognition of a system constitutes its criterion of validity: 'To say that a given rule is valid is to recognize it as passing all the tests provided by the rule of recognition and so as a *rule of the system*'.<sup>13</sup> The rule of recognition provides also the means for the resolution of conflicts between laws.<sup>14</sup> This is conceived by Hart as an essential part of the rule's function in identifying the laws of the system, for it is a condition of the validity of a rule that it does not conflict with a superior rule.

How does the rule of recognition fulfil its function? The rule is unique among the rules of the legal system.<sup>15</sup> It is a necessary rule in the sense that every legal system necessarily has one and only one rule of recognition and a set of rules which does not include a rule of recognition is not a legal system.<sup>16</sup> Furthermore,

---

if as a result of its recognition by the court in this case there arises a duty to apply it in other cases: that is, if the system has some rule of precedent.

<sup>9</sup> H. L. A. Hart, *The Concept of Law* (Oxford, 1961), pp. 92-120.

<sup>10</sup> The rule of recognition is also central to Hart's theory of the existence of legal systems. Cf. *ibid.*, pp. 109-13. This aspect of the doctrine will be disregarded in the present discussion.

<sup>11</sup> *Ibid.*, p. 92.

<sup>12</sup> *Ibid.* (emphasis added).

<sup>13</sup> *Ibid.*, p. 100 (emphasis added; cf. *ibid.*, pp. 102, 106).

<sup>14</sup> *Ibid.*, pp. 92-3, 98, 103.

<sup>15</sup> *Ibid.*, p. 107.

<sup>16</sup> Cf. *ibid.*, p. 93.

all the other laws exist and are part of the legal system only if they fulfil conditions laid down in the rule of recognition. The existence of the rule of recognition itself cannot, of course, be ascertained in this manner. 'Its existence is a matter of fact [and]... must consist in an actual practice.'<sup>17</sup> Hart offers a detailed analysis earlier in his book of what it means for a rule to exist as a matter of fact—as a social practice;<sup>18</sup> this analysis is quite clearly meant to apply to the rule of recognition.

Whose practice constitutes the conditions for the existence of the rule of recognition? Hart's answer is far from clear. Often he refers to 'the practice of courts, legislatures, officials or private citizens'.<sup>19</sup> On occasion, while including reference to the behaviour of private citizens, he attributes special importance to the practice of the courts.<sup>20</sup> Finally, we are told—and this should be regarded as Hart's position—that the behaviour of the population is not part of the conditions for the existence of the rule of recognition. Its existence consists in the behaviour of the 'officials' of the system,<sup>21</sup> by which he presumably means law-applying officials.<sup>22</sup>

Hart holds that the conditions for the existence of social rules are practices of those people to whom the rules are addressed. It follows, then, that the rule of recognition is addressed to the officials of the legal system.<sup>23</sup> Furthermore, Hart's explanation of social rules is basically an explanation of duty-imposing rules. The only other type of rules Hart recognizes are power-conferring rules, but he does not consider what social practices constitute the existence of a customary power-conferring rule.

<sup>17</sup> H. L. A. Hart, *The Concept of Law*, pp. 107–8; cf. p. 245.

<sup>18</sup> Cf. *ibid.*, pp. 54–6. For a discussion of Hart's explanation of a social rule see J. Raz, *op. cit.*, pp. 147 ff.

<sup>19</sup> H. L. A. Hart, *op. cit.*, p. 104 n. 9. See also *ibid.*, pp. 98, 106.

<sup>20</sup> *Ibid.*, p. 105.

<sup>21</sup> *Ibid.*, pp. 110, 113.

<sup>22</sup> This interpretation is confirmed by H. L. A. Hart, 'Kelsen's Doctrine of the Unity of Law', in *Ethics and Social Justice*, Kiefer and Munitz (eds.) (New York, 1970).

<sup>23</sup> There is no clear statement in the book on this issue. The rule is often said to be used by officials and individuals (cf. H. L. A. Hart, *The Concept of Law*, pp. 97, 98, 104), but it is quite clear that it is addressed to officials and applies only to them. It follows that if every legal system has a rule of recognition it also has officials and rules of change and adjudication investing officials with legal powers. That is the reason for my claim above that Hart's theory should be counted with the theories which emphasize the institutional nature of law. Occasionally, however, Hart tends to deny that institutions are an essential part of every legal system. Cf. *ibid.*, p. 93; Hart, 'Kelsen's Doctrine of the Unity of Law', p. 195.

Therefore, since the rule of recognition is a customary rule, it must be interpreted as duty-imposing. Besides, all the legal powers of officials are conferred on them by the rules of change and adjudication, authorizing them to make new laws and to settle disputes. To claim that the rule of recognition is a power-conferring rule is to confuse it with either rules of change or rules of adjudication.<sup>24</sup>

The rule of recognition imposes an obligation on the law-applying officials to recognize and apply all and only those laws satisfying certain criteria of validity spelt out in the rule, which criteria include indications of how conflicts of laws are to be resolved.<sup>25</sup> From Hart's examples it would seem that although he thinks that the criteria of validity most commonly refer to the mode of origin of the laws, this is not always the case.<sup>26</sup>

The jurisprudential criterion of identity implied by this theory is: A legal system consists of a rule of recognition and all the laws that ought to be applied according to it. When the courts apply a rule that they were not obliged to apply they may make it thereby into a law (if there is in the system a rule of precedent that will oblige all courts henceforth to apply it), but they do not apply an existing law and the rule was not part of the system before its application.

Hart's rule of recognition is subject to criticism on a few points and requires some clarification. One should remember that clear conceptual distinctions do not entail the existence of clear instances of the concepts involved. Therefore, the absence of clear instances should not deter one from striving

<sup>24</sup> It is commonly assumed that by secondary rules Hart means power-conferring rules. This interpretation is supported by some passages in his writings. Cf. Hart, Book Review, *Harvard Law Review* 78 (1965), 1281, 1292. This interpretation, however, conflicts with other aspects of his theory and does not represent his present views. It is true that all the primary rules are duty-imposing, but not all the secondary rules are power-conferring. The rule of recognition is an exception. Rules of change, adjudication, and recognition are called secondary because they presuppose the existence of primary rules, whereas primary rules can exist without secondary ones, albeit not as a legal system. This is an explanation of the terminology, not a criterion for determining which rules are primary and which are secondary. This is determined by their social function—whether they are rules of change, adjudication, and recognition—not by their normative character.

<sup>25</sup> The logical form of a rule of recognition is: All law-applying officials have a duty to apply all and only laws that satisfy the following criteria: ...

<sup>26</sup> Cf. H. L. A. Hart, *The Concept of Law*, pp. 92, 94, 97, 98.

to formulate clear conceptual distinctions. The courts, in most cases brought before them, probably neither merely apply an existing law nor do they merely initiate a new law. They may be doing a little of both. But this does not detract from the ability of a clear distinction between applying existing law and creating a new one to shed light on legal processes.

The application of the rule of recognition to concrete cases may be beset by similar problems. Hart makes it quite clear that the rule of recognition, like any other rule, is necessarily open-textured and vague to some extent. It may also be incomplete,<sup>27</sup> for it may not include an accepted answer to some legal questions, such as the validity of rules of public international law within the municipal legal system. The existence of a rule of recognition does not entail that all the legal problems, the solution of which may depend on the nature of the criteria of validity, such as the problems mentioned above, have found their solution in the system. So long as the rule is incomplete some such problems will remain unanswered, but when the courts are faced with such unsolved problems and accept a certain solution they modify the rule of recognition. This should surprise no one. The rule of recognition, being a customary rule, is constantly open to change.

Asserting that there is in every legal system a rule of recognition does not involve one in the task of giving a systematic and reasonable account of the limits of a legal system based on accepted or proposed solutions to a whole host of legal problems, as described above. However, any attempt to articulate the details of the criteria of validity incorporated in any rule of recognition of any legal system means embarking on precisely this endeavour. Attempting to formulate criteria of validity based on complex court practices that are in a constant state of change and that are necessarily vague and almost certainly incomplete, involves not only legal perceptiveness and theoretical skill, it demands sound judgment and reasonable value-decisions as well. Hart's theory leads one to the point where the boundaries between analytic and critical jurisprudence, between general and particular jurisprudence, begin to blur. But he himself does not cross the boundary. Rather, he provides

<sup>27</sup> H. L. A. Hart, *The Concept of Law*, pp. 144–50.

conceptual tools for dealing with particular and critical problems, but he does not deal with those problems himself.

There is no reason to suppose that the rule of recognition refers to all the criteria of validity of a legal system, and it is clearly wrong to think that it determines them all. A criterion of validity is a set of conditions set by law, satisfaction of which is sufficient for being a law of the system. All the laws of a legal system, except the rule of recognition the existence of which is a matter of social practice, are valid; they exist in the system because they satisfy some criterion of validity. Besides the rule of recognition, other laws can also set criteria of validity. All the laws conferring legislative powers, for example, determine criteria of validity; so also does a law stipulating conditions that a social custom must fulfil to be legally binding.<sup>28</sup>

That the rule of recognition sets up some criteria of validity is clear. There must be in every system some criteria of validity that, although legally binding, are not legally valid, hence they must be set in the rule of recognition. There is, however, no reason to think that all the criteria must be stipulated in that rule. The fact that all the criteria of validity are determined in laws that, directly or indirectly, are valid according to criteria determined in the rule of recognition, guarantees that by imposing a duty to apply the laws satisfying its criteria of validity, the rule of recognition imposes a duty to apply all the laws of the system.

Furthermore, there is no reason to suppose that every legal system has just one rule of recognition. It may have more. Imagine a legal system in which no valid law makes custom or precedent a source of law, but in which, nevertheless, both custom and precedent are sources of law. It follows that the criteria for the validity of laws created by custom or precedent are determined by rules of recognition imposing obligations on the courts to apply such laws. But we should not assume that there is just one rule of recognition rather than two—one relating to each source of law—simply because the

<sup>28</sup> Because it is a condition for being a law of a legal system that the courts ought to apply it, it follows that criteria of validity are conditions for the duty of the courts to apply laws. Therefore, since the courts have a duty to apply lawfully enacted laws, laws conferring powers of legislation give rise to criteria of validity. They are not, however, the only type of law that does so. The rule of recognition is just one example of a law setting criteria of validity that is not a power-conferring law.

system must contain means of resolving conflicts between laws of the various sources.<sup>29</sup> First, as was pointed out above, the rule of recognition, even if it is one rule, may be incomplete, which means that the system may not include any means of resolving conflicts. Perhaps the problem has never arisen and there is no generally accepted solution to it. Secondly, there may be two or more rules of recognition that provide methods of resolving conflicts; for example, the rule imposing an obligation to apply certain customs may indicate that it is supreme, whereas the rule relating to precedent may indicate that it is subordinate.<sup>30</sup>

In most legal systems courts have authority to settle at least some of the disputes to which there is no clear solution in the laws of the system. Courts have a duty to apply the laws of the system when they are applicable and to exercise discretion in order to decide (partially) unregulated disputes—disputes to which the laws do not provide a clear answer or where the courts have power to change the law. By the rule of precedent this exercise of discretion often amounts to the creation of new laws. The courts' discretion to decide unregulated disputes may be absolute or guided. They may be guided by law as to the manner in which discretion should be exercised. The law may, for example, direct judges to act on the rule that they think best for such cases, or to render a decision that would be the best from the point of view of the parties to the present dispute, or direct the courts to consult their conscience or the writings of moralists. Such instructions may be given in a statute, but they may also exist only in the practice of the courts. Very often, however, they are not ordinary precedents, but may instead derive their force from the continuing practice of the courts. When this happens, the legal system concerned contains two types of ultimate laws: laws of one type directing the courts which laws to apply, those of the other type guiding their discretion in deciding (partly) unregulated disputes. Laws of the first type are laws of recognition, laws of the second type are ultimate laws of discretion, and both impose duties on the courts. But laws of recognition oblige the courts to apply certain

<sup>29</sup> It is not clear whether Hart relies on this consideration, *op. cit.*, pp. 92–3.

<sup>30</sup> It is the fact that a set of laws of recognition are maintained by the practice of the same law-applying organs that indicates that they are all part of one legal system.

laws, leaving them no choice which laws to apply. Laws of discretion, on the other hand, whether ultimate or not, merely guide the courts' discretion in the choice of laws to adopt and apply; they limit the courts' freedom of choice but do not deprive them of it.

Laws of recognition are thus deprived of part of their uniqueness. They are still the only ultimate laws that necessarily exist in every legal system, but they are not the only ultimate laws that can exist in a system. Moreover, the distinction between applying an existing law and applying a new one is seen to be more a difference of degree than of kind. This fact, together with the fact that in practice it is often difficult to decide whether in a particular case a new law was created or an old one applied, does not mean that the distinction cannot be drawn or that it is unimportant. Every legal system rests on its ultimate laws, which commonly means on a set of ultimate laws of recognition and discretion. The former provide the ultimate criteria of validity of the laws of the system, the latter guide the courts in the exercise of their powers to modify the system when deciding unregulated disputes and creating precedents for the future. The difference may be one of degree, but it is indispensable for the formulation of criteria of identity.<sup>31</sup>

### **C. The Relation of Law and State**

If the theory of the rule of recognition is substantially correct, as I think it is, it forms part of the answer to the problem of identity. Although it sets necessary conditions for membership in a legal system, it does not provide all the sufficient conditions. Nothing is part of a legal system unless either it is a rule of recognition of the system, or the courts ought to recognize and apply it. To be a rule of recognition is sufficient to be counted as a law of the system, but to be a law that the courts are obliged to apply is not. Quite often the courts have an obligation to apply laws of other legal systems, rules of private associations, and so on, although these were not and do not become part of the legal system. Therefore, the rule of recognition provides

<sup>31</sup> In the next essay and in Essay 10 the test distinguishing between making and applying law is refined to explain how it is that a court is under an obligation to apply a law which it can, under certain conditions, revise.

no complete answer to the problem of the scope of a legal system—the problem of the identity of momentary legal systems.<sup>32</sup>

Nor does the rule of recognition solve the problem of the continuity of legal systems. That one legal system comes to an end and another takes its place manifests itself in a change of rule of recognition, for each legal system has a different rule of recognition. The rule of recognition, however, is a customary rule; hence it is constantly in a process of change. What changes are consistent with the continued existence of the same rule, and what changes compel the admission that a new rule has replaced the old one?<sup>33</sup> It is easy to bring examples for either situation, as well as examples of borderline cases. However, it is not the existence of borderline cases, which are inevitable, that is disturbing. The disturbing fact is that Hart's theory provides no clue as to *how* to draw the conceptual distinction. Even more disturbing is that this is no mere oversight on the part of Hart that can be easily remedied. He did not provide the answer because he did not ask the question to which the distinction is an answer. His theory provides substantially complete answers (whether or not they are correct) to the problems with which he was concerned: the role of the courts in a legal system, the truth in rule scepticism, the variety of laws and their interrelation, the relation of efficacy and existence. If his theory fails to provide a complete solution to the problem of identity it is because he overlooked not only part of the answer but also a whole question: that of the relation of law and state.<sup>34</sup>

The relation of law and state affects the two distinct aspects—scope and continuity—of the problem of identity. Every state—by which is meant a form of political system and not a juristic person—has one legal system that constitutes the law of that

<sup>32</sup> Hart has become aware of these problems. Cf. H. L. A. Hart, 'Kelsen's Doctrine of the Unity of Law', *op. cit.*, pp. 195 f. The fact that a legal system may contain more than one rule of recognition is another reason for the need to revise Hart's doctrine of identity.

<sup>33</sup> See generally J. Finnis, 'Revolution and Continuity of Law' in A. W. B. Simpson, *Oxford Essays in Jurisprudence*, 2nd series (Oxford, 1973).

<sup>34</sup> In many legal systems the term 'state' or its equivalent is used to designate a certain legal person recognized in law as having certain duties, powers, and rights, and as acting through certain organs. This legal concept of the state should be clearly distinguished from the political concept of the state as a form of political system; only the latter is relevant to the problem of identity.

state, and every municipal legal system is the law of one state. Since, then, the identity of a legal system is bound up with that of the state the law of which it is, the relation between law and state necessarily affects the problem of scope. So too, since an end to the existence of a state is the end of its legal system, and since a law that is not a law of the state is not part of its legal system, the problem of continuity is similarly affected by the relation of state and law.

Two diametrically opposed views on the relation of law and state have been expressed by legal philosophers. Kelsen<sup>35</sup> claimed that the concept of the state can be explained only in legal terms. That is, the concept of a legal system must be explained first; from it naturally flows the explanation of the concept of a state, for a state is but a (municipal) legal system. No social facts, no social norms that are not relevant to the explanation of law have any relevance to the theory of the state. Bentham and Austin, on the other hand, held that law can only be explained after some theory of state has been established. First one must define the meaning of 'an independent political society'—'a state'. On the basis of this definition 'law' can be defined. Bentham's and Austin's definitions of 'an independent political society' are purely sociological, making use of no legal concepts, and the same is true of their definitions of sovereign and subject that are part of it.<sup>36</sup>

Bentham and Austin have the better of this controversy. Because Kelsen lacks the concept of the state as a political system, he fails to account for the identity of a legal system. He is driven to rely on constitutional continuity as a sole mark of identity, disregarding the fact that new states can be created and new legal systems established without any break in the constitutional continuity taking place.<sup>37</sup> A theory of law must be based, at least partly, on a theory of state, and denying this has been one of Kelsen's gravest mistakes. A theory of state, however, is partly based on a theory of law—the two are intimately interrelated.

Since I am concerned with the nature of the problem of

<sup>35</sup> H. Kelsen, *General Theory of Law and State*, pp. 181–207.

<sup>36</sup> Cf. J. Austin, *The Province of Jurisprudence Determined* (New York, 1954), pp. 192–215; J. Bentham, *A Fragment of Government* (Oxford, 1960), pp. 33–46.

<sup>37</sup> For a detailed criticism of Kelsen's theory of the identity of legal systems see J. Raz, *The Concept of a Legal System*, pp. 95–109.

identity rather than its solution, there is no need to discuss the concept of a state beyond mentioning some truisms. A state is the political organization of a society, it is a political system that is a subsystem of a more comprehensive social system. The social system includes, of course, many other subsystems and the political system interacts with most if not all of them, as well as interacting with other political systems. These social and political systems are normative systems; in other words, at least part of the pattern of interrelations constituting the systems is norm governed.

The legal system is only part of the norms constituting the political system; most political systems include numerous non-legal norms. Some of these non-legal norms apply to society in general: however much the system is resented and hated, if it is viable at all, it is based, among other things, on some norms of respect for at least some of the laws and some authorities on the part of some important sections in the society. Some non-legal political norms are more limited, of which Dicey's conventions of the constitution can serve as an example.<sup>38</sup>

It follows that since the continuity of a legal system is tied to the continuity of the political system, the former is affected by the fate of the non-legal norms that happen to form part of the political system concerned. However, emphasizing the importance of the fate of non-legal norms to the continuity of the legal system does not mean that these are the only factors affecting continuity. The substance of my contention is that whatever form one's ultimate account of continuity<sup>39</sup> takes, it must, in view of the relation of law and state, be based on the following two points: first, that continuity depends on the interaction of legal and non-legal norms, and the extent and manner of their change; and secondly, that among the legal norms concerned some are more relevant than others. Since the continuity of the legal system is fundamentally a function of the continuity of the political system, political laws are more relevant than others. Constitutional and administrative laws are, therefore, more relevant than, for example, the law of contract or torts.

<sup>38</sup> Sec A. V. Dicey, *Introduction to the Study of the Law of the Constitution* (London, 1963).

<sup>39</sup> It may, for example, be based on the conception of change of system as a result of a short and intense period of fundamental change, or it may allow for a slow and gradual transformation of one system into another.

The problem of scope is similarly affected, and can be subdivided into four sub-problems: first, that of the dividing line between political norms that are part of the legal system and those that are not; secondly, that of the dividing line between legal norms and social norms of the social system of which the political system is a subsystem; thirdly, that of the dividing line between the law and norms of other subsystems of the same society; and fourthly, that of the dividing line between one legal system and coexisting laws of other legal systems.

The first of these sub-problems of the problem of scope may be solved by insisting that the laws of a system are either the ultimate rules of its courts or the laws its courts ought to recognize and apply. Those political norms that are neither the courts' practice nor norms that the courts ought to apply are not part of the law of the state. The other three sub-problems, though each involving somewhat different considerations, have this much in common: sometimes the courts are under an obligation to apply norms because those norms belong to these other social or political systems. The courts ought to enforce private contracts, the rules of some private associations within the state, the laws of foreign countries, and so on. Some theorists have seen this as a reason for regarding those laws as part of the legal system, and Bentham maintained that all commands that are enforceable in law are laws of the sovereign.<sup>40</sup> Hart maintained in *The Concept of Law* that all the rules that the courts have a duty to apply are laws of the system.

As indicated above, the backing of the state power is a defining characteristic of municipal law,<sup>41</sup> but it is not the only one. Also characteristic of the state system is that one of its main functions is to maintain and support other forms of social grouping; it is, therefore, characteristic of the law that it upholds and enforces contracts, agreements, rules, and customs of private persons and associations. To obscure the distinction between norms recognized as part of the law and norms that, although not part of the law, are recognized and enforced because it is the function of the law to support various social groupings is

<sup>40</sup> J. Bentham, *Of Laws in General* (London, 1970), pp. 22 ff.

<sup>41</sup> I have limited the discussion throughout to municipal law. Other types of legal systems are the law of other types of social organizations, be they tribes, churches, or the international community and they bear similar relations to those organizations.

to misunderstand the nature of the state and its relations to other social systems.

Admitting that not all the norms that the courts ought to apply are part of the law, where should one draw the line? This is not the place to attempt a solution.<sup>42</sup> I wish, however, to conclude with one last remark bearing on the problem. That a norm is identified as one the courts ought to apply by the fact that it is a norm of a certain society, association, or state is no indication whether or not it is part of the system. Legislation by reference is a familiar technique; for example, a statute passed in one country adopting by reference the civil code of another country. No other formal distinction will succeed in drawing a reasonable dividing line. The reasons for enforcing the norm, and the attitude of the courts and the legislature to its enforcement, are the crucial factors. Formal distinctions may give some indication as to the nature of the reasons for enforcing, but are never in themselves conclusive. Ultimately the problem turns on an accumulation of evidence justifying a judgment whether the norm is enforced on the grounds that it is part of the law's function to support other social systems or because it is part of the law itself.

<sup>42</sup> Cf. Essay 6 below.

# The Institutional Nature of Law\*

## I. Introduction

My purpose is to outline one view concerning the identifying features of municipal legal systems. Analytical jurisprudence deals with three main problem areas. One concerns the special features of the judicial process and of judicial reasoning. The second encompasses the discussion of legal concepts (e.g. rights, duties, ownership, legal person) and of types of legal standards (rules and principles, duty-imposing standards, and power-conferring standards). The third range of problems revolves round the idea of a legal system and the features which distinguish such systems from other normative systems. It is with some of the problems belonging to this area that I shall be concerned.

It may be useful to begin by confessing to some of my prejudices and to make explicit a few of the assumptions underlying my reflections. One of them can be dubbed the assumption of the primacy of the social. We are familiar with the distinction between legal systems which are in force in a certain society and those which are not. There is a legal system in force now in Great Britain and there is one in force in Norway. But the legal system once in force in the Roman Republic is no longer in force, nor is the legal system proposed by a group of scholars for country X in fact in force in that country. Whether or not a system is in force in a society depends on its impact on the behaviour of people in the society. The precise nature of the criterion determining if a system is in force is a disputed issue with which I will not be concerned. But whatever it is, it concerns the attitudes and responses of all or certain sections in the society to the legal system: Do they know it, do they respect it, obey it?, etc. This seems to me to be a very significant fact. The reason is that whether or not a system is in force is not just another question about a legal system comparable with questions such as: Is it a Socialist or a Capitalist legal system?

\* First published in *The Modern Law Review* (1975). The argument of this essay is further defended in sections 4.2, 4.3, 5.1, and 5.2 of *Practical Reason and Norms* (2nd ed., Oxford, 1999), which incorporate part of it.

Is it a Federal system or not?, etc. We identify the question of whether or not the system is in force with the question of its existence. A legal system exists if and only if it is in force. The significance of the point is that it brings out that normative systems are existing legal systems because of their impact on the behaviour of individuals, because of their role in the organization of social life. Consequently when we look at legal systems as systems of laws, when we consider their content and disregard the question of whether they are in fact in force, whether they exist, we should look for those features which enable them to fulfil a distinctive role in society. These will be the features which distinguish legal systems from other normative systems. This is the assumption of the primacy of the social. It does not mean that legal systems do not also have other characteristic features. They may, for example, have certain moral features. It may be a necessary truth that all legal systems conform to some moral values and that a system which violates those values cannot be a legal system. My claim is merely that if this is indeed the case then these necessary moral features of law are derivative characteristics of law. If all legal systems necessarily possess certain moral characteristics they possess them as a result of the fact that they have other properties which are necessary for them to fulfil their unique social role.

There are two other assumptions which I wish very briefly to mention. The first is the assumption of universality according to which it is a criterion of adequacy of a legal theory that it is true of all the intuitively clear instances of municipal legal systems. Since a legal theory must be true of all legal systems the identifying features by which it characterizes them must of necessity be very general and abstract. It must disregard those functions which some legal systems fulfil in some societies because of the special social, economic, or cultural conditions of those societies. It must fasten only on those features of legal systems which they must possess regardless of the special circumstances of the societies in which they are in force. This is the difference between legal philosophy and sociology of law. The latter is concerned with the contingent and with the particular, the former with the necessary and the universal. Sociology of law provides a wealth of detailed information and analysis of the functions of law in some particular societies. Legal

philosophy has to be content with those few features which all legal systems necessarily possess.

My third assumption has already been incorporated into my second criterion of adequacy by restricting it to municipal legal systems. It can be called the assumption of the importance of municipal law. It reflects our, or at least my, intuitive perception that municipal legal systems are sufficiently important and sufficiently different from most other normative systems to deserve being studied for their own sake. They are, or are part of, a form of social organization which is both important and different from most others and which therefore should be made an object of a separate study. Obviously, in part the investigation of municipal systems is designed to compare and contrast them with other normative systems. Indeed it is to this part that the present essay is dedicated. In pursuing such investigations it may turn out that municipal systems are not unique, that all their essential features are shared by, say, international law or by church law. If this is indeed so, well and good. But it is not a requirement of adequacy of a legal theory that it should be so or indeed that it should not be so. It is, however, a criterion of adequacy that the theory will successfully illuminate the nature of municipal systems.

## **II. Primary Institutions**

Many, if not all, legal philosophers have been agreed that one of the defining features of law is that it is an institutionalized normative system. Two types of institutions were singled out for special attention: norm-applying institutions such as courts, tribunals, the police, etc., and norm-creating institutions such as constitutional assemblies, parliaments, etc. I have argued elsewhere that the existence of norm-creating institutions though characteristic of modern legal systems is not a necessary feature of all legal systems, but that the existence of certain types of norm-applying institutions is. Let me try to say something about the nature of institutions which are a necessary part of every legal system. What are the distinguishing marks of norm-applying organs? This is a notoriously difficult question. We have only to look at the debate concerning the nature of courts to become aware of the difficulties. Lawyers and sociologists have offered various incompatible explanations and

the battle is still raging. Given this history of disagreement the first thing to note is that various theorists studying this question are really tackling a variety of problems. Lawyers studying the defining features of 'a court' or 'a tribunal' may be concerned with solving any one of a variety of legal problems arising under a specific legal system: a certain court in that system may have supervisory powers over all judicial determinations by judicial bodies. The law of evidence or some of its rules may apply to proceedings before every judicial body, etc. When a lawyer faces the question: 'What is a court?' he is usually concerned with one or more of the many problems which such laws give rise to. Is the body A a judicial body subject to the supervisory jurisdiction of the relevant court? Do the general principles of the law of evidence apply to proceedings before A?, etc. Social scientists have their own problems which are quite different, though usually indirectly related to those of the lawyer. They may be interested in the classification of different social methods of settling disputes, or the different channels for the articulation of demands, etc. Our purpose in looking for the identifying traits of norm-applying institutions is primarily to establish the nature of the kind of institutions the presence of which is a defining feature of legal systems. An adequate answer to our question need not be a satisfactory solution of the problems of the lawyer or of the sociologist, nor is it intended as an answer to their questions.

Some have attempted to define judicial and other norm-applying organs by the social functions they fulfil. Others have looked for an answer in the norms which establish these institutions. I shall follow the latter kind of approach. Norm-applying institutions are first and foremost normative institutions established by norms and it is to these we must turn for a clue to their identity. It may be true that they are established to serve some social functions, but it is likely that the same functions can be and are served by other means as well. Norm-applying institutions should, therefore, be identified by the way they fulfil their functions rather than by their functions themselves. This does not detract from the importance of studying the functions which the institutions serve, it merely means that the institutions have to be identified by other means.

Of legal systems it can be said that every act by a public

official which is the performance of a duty or an exercise of a regulative power is generally regarded as a law-applying act. A policeman arresting a suspect, an official granting a trader's licence, a court rendering judgment in which Doe is ordered to pay a sum of money to Roe—all these are commonly regarded as instances of the application of law by public officials. These cases differ from similar acts of private individuals who pay taxes, make contracts, give orders to their employees, etc., only in being the acts of public officials. Therefore, on the most general interpretation of 'norm-applying institutions' these are identical with public institutions (in one sense of the word 'public').

What are the identifying features of public officials? This is a problem which is both important and difficult. It is, however, a problem which it would be best to avoid here, for though we will find public officials in all legal systems, not all of them must exist in the system if it is to count as a legal system. Instead we should try to identify a subclass of norm-applying institutions, namely those the presence of which is necessary in all legal systems.

The terminological contrast between 'norm-creating' and 'norm-applying' draws attention to one important class of norm-applying institutions—those which apply norms not by making other norms but by physically implementing them. The courts apply the law by rendering judgments which are themselves norms. The prison service or public officials instructed to pull down a house against which a demolition order has been issued physically enforce the law. I shall call norm-applying institutions of this kind norm-enforcing institutions. There is no doubt that norm-enforcing institutions play an important role in all modern legal systems. Yet they cannot be regarded as the key to the identification of legal systems. Though all legal systems regulate the use of force and ultimately rely on force to ensure compliance with the law, not all of them need have law-enforcing institutions. There may be normative systems which share all the characteristics of legal systems and do not have law-enforcing machinery. Once a judgment is given its execution is left to the parties to the dispute. In such a system an individual is not allowed to use force to secure his rights whenever he likes. He is obliged to go to a court and

obtain an authoritative declaration of his rights. But once he is in possession of a decision he is entitled to implement it using reasonable force and he may be entitled to authorize others to use force in his name for this purpose. Such a system is clearly a legal system. It does not have law-enforcing institutions, but it has other norm-applying institutions which warrant regarding it as a legal system.

We must, therefore, look elsewhere for the kind of norm-applying institutions which are crucial to our understanding of legal systems. I shall suggest that the type of institutions we are looking for are those which combine norm-making and norm-applying in a special way. Let us call these institutions primary (norm-applying) organs, to indicate their importance. Primary institutions are just one kind of norm-applying institution. Norm-enforcing organs are another kind of such institution and there are others as well. Norm-enforcing organs are concerned with the physical implementation of norms and this determines their character as norm-applying. Primary organs are concerned with the authoritative determination of normative situations in accordance with pre-existing norms. Consider judicial bodies. Courts and tribunals have power to determine the rights and duties of individuals. But cannot any person do the same? Cannot John determine whether he owes £100 to Alan or whether Paul owes money to Jack? He may be ignorant of the facts but like a court he may investigate them. He may be ignorant of the law but like a court he may study it. The difference between a court and a private individual is not merely that courts are provided with better facilities to determine the facts of the case and the law applying to them. Courts have power to make an *authoritative* determination of people's legal situation. Private individuals may express their opinion on the subject but their views are not binding.

The fact that a court may make a binding decision does not mean that it cannot err. It means that its decision is binding even if it is mistaken. My declaration of the legal situation is not binding at all because it is not binding if it is mistaken. To be a binding application of a norm means to be binding even if wrong, even if it is in fact a misapplication of the norm. This seemingly paradoxical formulation illuminates the nature and function of primary norm-applying organs.

The paradox is generated by the following problem: How can we say of a determination (decision or declaration) both that it applies a pre-existing norm and that it is binding? We may feel that we regard a determination as norm-applying if it merely determines which rights and duties individuals have in virtue of pre-existing norms, whereas we regard a determination as binding only if it changes the rights and duties of individuals. Only with regard to a new norm imposing duties on individuals or releasing them from their duties, investing them with rights or divesting them of their rights can we ask whether it is valid or not. If the determination purports merely to ascertain what rights and duties they already have and not to change them then the only question arising is whether the determination is correct or incorrect. The question of the binding force arises only with respect to creative determinations—those which change the normative situation. Creative determinations can be binding or not but cannot be either correct or incorrect. The reverse is true of applicative determinations.

On this view a determination cannot be both binding and norm-applying. This is, however, an over-restrictive view of the sense of 'binding'. A determination can be binding even if it does not change the normative situation, provided it would have been binding had it changed it. Consider a new piece of legislation which, though its authors may be unaware of this fact, merely repeats the content of an old but valid law. The new legislation can be judged to be either valid or invalid, even though it is clear that it does not change anybody's rights or duties. The point is that if valid it would have changed the legal situation had the old law no longer been in force. Put in another way—if valid it creates another basis for the rights and duties imposed by the old law. In the same sense a court's determination that Doe owes money to Roe is binding even though the debt existed by virtue of a pre-existing norm, provided that it is binding even if there would have been no debt but for the decision of the court—hence my original formulation that a norm-applying determination is binding only if binding even if mistaken.

Now we are in a position to describe the defining features of primary norm-applying organs: they are institutions with power to determine the normative situation of specified

individuals, which are required to exercise these powers by applying existing norms, but whose decisions are binding even when wrong. A few comments on this characterization will be in place here:

(1) The definition attempts to identify one kind of institution. The nature of institutions in general is presupposed and is not explained in it. It is important to stress that we are concerned with primary *institutions*. Legal systems are not identified merely by the fact that they contain norms conferring powers to make binding applicative determinations. They must contain norms conferring such powers on institutions, i.e. on centralized bodies concentrating in their hands the authority to make binding applicative determinations.

(2) Courts, tribunals, and other judicial bodies are the most important example of primary organs. But other officials, such as police officers, may also be primary organs. There are obvious reasons to impose on primary organs a duty to follow judicial procedures, but this need not be always done. It does seem reasonable to suppose, however, that the notion of a primary institution provides a necessary step in any attempt to analyse the nature of judicial institutions.

(3) The definition of a primary organ may have to be further refined. As it stands it applies only to final and absolutely binding determinations. It has to be modified to allow for the possibility of appeal, re-trial, etc., and also for the possibility that the determination is binding for one purpose but not for others. In many legal systems there are applicative determinations which are binding only with respect to the cause of action whose litigation resulted in the determinations.

(4) The definition identifies primary organs by their power to make binding applicative determinations. This is compatible with the fact that the same institutions have other powers and functions. In particular, courts often have power to create precedent and lay down general rules, to issue orders to individuals to perform certain actions, and authoritatively to determine the facts of the case (the *res judicata* doctrine). All these are either entirely different or, at best, overlap with the power to make binding applicative determinations. Applicative determinations are determinations of the rights or duties of individuals in concrete situations and are entirely different from the power

to create precedent or to issue orders instructing individuals to pay damages or fines or be gaoled, etc., because they disregarded their duties or the rights of others. Applicative determinations are most closely related to declaratory judgments. In fact the definition suggests that a declaratory judgment is an ingredient in many courts' decisions. This is part of the effect of the *res judicata* doctrine. But this doctrine is wider and applies also to purely factual findings, and not only to determinations of rights and duties in particular situations.

### III. The Limits of Law

My claim is that our common knowledge of intuitively clear instances of municipal systems confirms that they all contain primary institutions and that such institutions play a crucial role in our understanding of legal systems and their function in society. The analysis both of the existence conditions and of the identity of legal systems depends on the rules governing the working of their primary institutions and their actual behaviour. Furthermore, it seems reasonable to suppose that the law differs from any other methods of social control in providing machinery for the authoritative settlement of disputes. To explore these issues will take us beyond the scope of the present essay.

I would like, however, to point to one consequence of the necessary existence of primary institutions. We saw that their presence means that law provides a method for settling disputes. It is important to notice that this is a special method for settling disputes. Consider a normative system containing only rules instituting tribunals and regulating their operation. When a dispute is referred to a tribunal it will be authoritatively settled by its decision, for the system includes a rule to that effect. It contains a rule making it obligatory to observe the decisions of the courts and providing for their enforcement by a police force created for this purpose only. In such a system the courts can settle any dispute in any way they like. There are no legislated, customary, or any other standards which they have to apply. Nor do they have to follow their own precedents. The courts of this peculiar system are not entitled to decide in an arbitrary way. They are instructed by a rule of the system to make that decision which seems to them the best in the

circumstances, taking account of all the considerations which seem to them relevant. Since the courts have to act on reasons and reasons are general, we could expect some regularity in the decisions of the courts. The same judge hearing two very similar cases on the same day is likely to reach the same decision in both. But cases will be heard by many different judges and judges may change their mind as well as forget complex and intricate arguments, etc. Consequently, the decisions of the courts over a period of time are unlikely to reflect any consistent line on any one issue. The degree of regularity will depend on contingent factors such as the number of judges, the degree of uniformity of their social background, etc.

It is unlikely that such a system has ever existed or will ever exist. I am conjuring up the image of a system of this nature simply because it resembles legal systems in having courts with powers to settle disputes. Therefore, by contrasting it with legal systems we can better observe what other features legal systems necessarily possess. Since its judges are not obliged to follow any common standards and can decide whatever they think best, the system does not provide any guidance to individuals as to how to behave in order to be entitled to a decision in their favour, should a dispute arise. Legal systems, on the other hand, do provide guidance to individuals. They contain laws determining the rights and duties of individuals. These are laws which the courts are bound to apply in settling disputes and it is because of this that they also provide an indication to individuals as to their rights and duties in litigation before the court.

Am I merely stating the obvious, namely, that legal systems must include laws, including some which are addressed to the general population? I think that two further consequences are contained in what I said, consequences which are far from being trivial. In the first place, law contains both norms guiding behaviour and institutions for evaluating and judging behaviour. The evaluation is based on the very same norms which guide behaviour. Indeed the test by which we determine whether a norm belongs to the system is, roughly speaking, that it is a norm which the courts ought to apply when judging and evaluating behaviour.<sup>1</sup> Thus the law can be said to possess its own internal system of evaluation. We can assess behaviour from

<sup>1</sup> One modification of this test is introduced in the next section.