

*Foundations of Social Theory*

JAMES S. COLEMAN

*The Belknap Press of Harvard University Press  
Cambridge, Massachusetts, and London, England*

1990

6. In an escape panic the more prominent an individual's position in the crowd (the greater the attention directed to him), the more likely he will be to exhibit orderly exit behavior.
7. The larger a crowd is in absolute numbers, the more likely it is to break out into hostile or expressive collective action.
8. The larger a crowd is in absolute numbers, the greater is the likelihood that a panic will occur in a physical escape situation.
9. Such a relation will not hold for bank panics.
10. Physical escape panics (sequence of actions) should show greater variability in outcome when circumstances are similar than is true for bank panics (single action).
11. Contagious beliefs should arise at times of extensive social change, when rights of control have been withdrawn from institutions that have power.

« 10 »

### *The Demand for Effective Norms*

Much sociological theory takes social norms as given and proceeds to examine individual behavior or the behavior of social systems when norms exist. Yet to do this without raising at some point the question of why and how norms come into existence is to forsake the more important sociological problem in order to address the less important. Whatever the reason for neglect of this question (and it differs for different theorists), I will show in this chapter and the next that two simple conditions, taken together, are sufficient for the emergence of norms. The first of these, to be examined in this chapter, is a condition under which a demand for effective norms will arise. The second, to be examined in the next chapter, is a condition under which that demand will be satisfied. Both conditions may be described as social-structural.

As much as any other concept in the social sciences, a norm is a property of a social system, not of an actor within it. It is a concept that has come to play an extensive role in theories developed by some sociologists. The reasons are fundamental. The concept of a norm, existing at a macrosocial level and governing the behavior of individuals at a microsocial level, provides a convenient device for explaining individual behavior, taking the social system as given. This device has been especially useful for those sociologists characterized by Sorokin (1928) as members of the sociological school of social theorists, of which Emile Durkheim was the most prominent member. Durkheim began with social organization and in a part of his work asked, "How is an individual's behavior affected by the social system within which he finds himself?" Answering this requires not the three components of social theory that I outlined in the first chapter of this book but only one—the transition from macro to micro. For many social theorists, Durkheim among them, the concept of a norm provides a means for making this transition.

For another school of social theory, of which Talcott Parsons is the most prominent member, the concept of a norm provides a basis for a principle of action whose role in the theory is comparable to that of maximizing utility in rational choice theory. The principle, something like "Persons behave in accordance with social norms," leaves examination of the content of norms as the theoretical task at the macro level. Whereas rational choice theory takes individ-

ual interests as given and attempts to account for the functioning of social systems, normative theory takes social norms as given and attempts to account for individual behavior.

Apart from its role in social theory, the use of the concept of a norm is important in describing how societies function. This is especially so for the description of traditional stable societies. A description of the functioning of the caste system in India that did not use the concept of dharma, which means something like "duty" or "appropriate behavior" or "behavior in accordance with accepted norms," would hardly be possible.<sup>1</sup> Stable or slowly changing norms constitute an important component of a stable society's self-governing mechanisms.

Both the evident importance of norms in the functioning of societies and the importance of a norm as a concept throughout the history of social theory underlie the importance of this concept in contemporary social theory. It has not one but two entries in the *Encyclopedia of the Social Sciences* (both written by sociologists), and one of them begins with this sentence: "No concept is invoked more often by social scientists in explanations of human behavior than 'norm.'" For example, Dahrendorf (by no means one of those sociologists most wedded to the concept), in an essay on the origin of social inequality, states, "The origin of inequality is thus to be found in the existence in all human societies of norms of behavior to which sanctions are attached . . . the derivation suggested here has the advantage of leading back to presuppositions (the existence of norms and the necessity of sanctions) which at least in the context of social theory may be taken as axiomatic" (1968, p. 104).

Norms may be taken as axiomatic by many sociologists, but for others they constitute an unacceptable *deus ex machina*—a concept brought in at the macrosocial level to explain social behavior, yet itself left unexplained.

Some rational choice theorists, armed with maximization of utility as a principle of action, regard the concept of a norm as altogether unnecessary. To take this stance, however, is to ignore important processes in the functioning of social systems and thus to limit the theory. It is one thing to refuse to take norms as starting points for social theory; it is quite another to ignore their existence altogether. In this book I refuse to take norms as given; in this chapter I ask how norms can emerge and be maintained among a set of rational individuals.

Social norms enter the theory developed here in the following way: They specify what actions are regarded by a set of persons as proper or correct, or improper or incorrect. They are purposively generated, in that those persons who initiate or help maintain a norm see themselves as benefiting from its being observed or harmed by its being violated. Norms are ordinarily enforced by sanctions, which are either rewards for carrying out those actions regarded as correct or punishments for carrying out those actions regarded as incorrect.

1. See O'Flaherty and Derrett (1978).

Those subscribing to a norm, or, as I will say, those holding a norm, claim a right to apply sanctions and recognize the right of others holding the norm to do so. Persons whose actions are subject to norms (who themselves may or may not hold the norm) take into account the norms, and the accompanying potential rewards or punishments, not as absolute determinants of their actions, but as elements which affect their decisions about what actions it will be in their interests to carry out.

In the preceding paragraph I do not give an explicit definition of a norm but only indicate its function. The explicit definition is, however, important because it derives from the conception of rights discussed in Chapter 3 and because it may not include everything that is ordinarily meant by the concept of a norm. I will say that a norm concerning a specific action exists when the socially defined right to control the action is held not by the actor but by others. As discussed in Chapter 3, this implies that there is a consensus in the social system or subsystem that the right to control the action is held by others. By the definition of authority, this means that others have authority over the action, authority that is not voluntarily vested in them, either unilaterally or as part of an exchange, but is created by the social consensus that placed the right in their hands. The right that is relevant to the definition of a norm is not a legally defined right or a right based on a formal rule imposed by an actor having authority. It is, rather, an informal or socially defined right. It may exist in the absence of a legally defined right or in opposition to a legally defined right, as is the case when a norm is in conflict with a law.

With this definition, the question concerning the conditions under which an effective norm will arise becomes a question concerning the conditions under which there will come to be a consensus that the right to control an action is held by persons other than the actor, and the conditions under which that consensus can be enforced.

It is important to note that this definition is a very specific and perhaps narrow one. No norm exists as long as the individual actor holds the right to his own action, and no norm exists if no right has come into existence. A norm exists only when others assume the right to affect the direction an actor's action will take. But *when* does this happen? This chapter is aimed at answering this question. There are, however, additional questions to be answered if norms are to be fully understood.

A norm may be embedded in a social system in a more fundamental way: The norm may be internal to the individual carrying out the action, with sanctions applied by that individual to his own actions. In such a case a norm is said to be internalized. An individual feels internally generated rewards for performing actions that are proper according to an internalized norm or feels internally generated punishments for performing actions that are improper according to an internalized norm. How and when does this happen?

There is interdependence among norms such that many norms are part of a

structure of norms. The most elaborate of such structures are those described by dharma in India and analogous systems in other societies with long cultural traditions. How do these structures come into being?

These questions pose substantial tasks for the theorist. First is the task of establishing the conditions under which a norm with a particular content will arise. This includes determining why a norm does not always arise when the existence of an effective norm would be in the interests of all or most persons. Related to this are the tasks of specifying who will come to hold the norm and whose actions will be the target of the norm. Another task is determining the strength and prevalence of sanctions, recognizing that applying a sanction may entail costs for the sanctioner. Related to this is determining what kinds of sanctions will be applied, since there are a variety of sanctions that may be applied (and it is empirically evident that various kinds of sanctions are applied, ranging from those that damage or enhance reputations to those that impose physical damage or provide material benefits). In addition, there are theoretical tasks concerning the internalization of a norm. Why do persons attempt to induce internalization in others in the first place? Under what conditions will those who hold a norm attempt to induce internalization, and under what conditions will they use only external sanctions? Why will a person be receptive to attempts by others to internalize norms? Finally, there is the task of describing and accounting for interconnections among norms. What kinds of relationships exist among norms, how do those relationships arise, and how is the role that norms play in a social system affected by these relationships?

It is useful to begin by locating the concept of a norm, as well as the theoretical activity of this chapter, in the context of the three components which I have proposed as necessary to social theory: the macro-to-micro transition, purposive action at the micro level, and the micro-to-macro transition. Norms are macro-level constructs, based on purposive actions at the micro level but coming into existence under certain conditions through a micro-to-macro transition. Once in existence, they lead, under certain conditions, to actions of individuals (that is, sanctions or threat of sanctions) which affect the utilities and thus the actions of the individuals to whom the sanctions have been or might be applied. Thus norms constitute a social construction which is a part of a feedback process, involving either negative feedback, which if effective discourages and dampens certain actions, or positive feedback, which if effective further encourages certain actions.

The emergence of norms is in some respects a prototypical micro-to-macro transition, because the process must arise from individual actions yet a norm itself is a system-level property which affects the further actions of individuals, both the sanctions applied by individuals who hold the norm and the actions in conformity with the norm. A diagram analogous to Figure 1.2 but illustrating the emergence of a norm begins at the micro level and ends there as well, with individual sanctions and conformity to the norm, as shown in Figure 10.1.

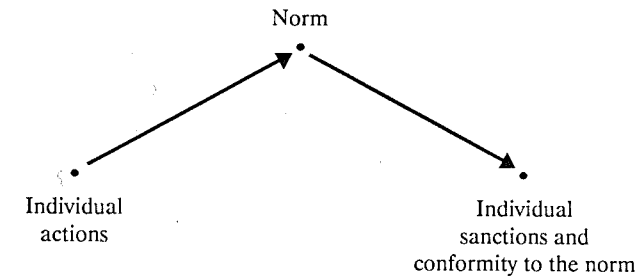


Figure 10.1 Relations of micro and macro levels in the emergence of a norm.

### Examples of Norms and Sanctions

Some sense of what is meant by norms and sanctions can be gained by considering several examples.

1. A three-year-old child, walking with its mother on a sidewalk in Berlin, unwraps a small piece of candy and drops the cellophane on the sidewalk. An older woman who is passing by scolds the child for dropping the cellophane and admonishes the mother for not disciplining the child. A three-year-old child, walking with its mother on a sidewalk in New York City, unwraps a piece of cellophane and drops the paper on the sidewalk. An older woman is passing by but says nothing, not even noticing the action of the child. Several questions are raised by this example: Why does the woman in Berlin assume the right to scold the child and admonish the mother? Why does a woman in a similar circumstance in New York City not do the same? Does the woman in New York not feel she has the right to scold the child, or does her failure to act arise from other sources?

2. In an organization which provides free coffee and tea to its employees, one employee who drinks tea goes down with his cup to the hot water dispenser. All the tea bags are gone, but he expresses no dismay, remarking to another person standing there, "This often happens, but I have taken some tea bags back to my office just for such occasions." The other person responds in a disapproving way, "It's people like you, stashing tea bags away, who create the problem." This example also raises questions: How did the second person come to acquire a right to express disapproval? And why did the first person leave himself open for such a comment, by his remarks? Furthermore, why does he accept the disapproval of the second person, apparently acknowledging the right of the second person to impose this sanction?

3. A high-school girl on a date at a beach house finds herself in a crowd in which the others, including her date, are smoking marijuana. The others encourage her to do so as well, showing disapproval and disdain of her reluctance. That reluctance, in turn, is produced by her knowledge that her parents would disapprove. This example raises questions about conflict: Can there be two conflicting

norms governing the same action? If so, then what determines which one, if either, will govern? And if conflicting norms do occur, in what class of situations do they arise?

4. Among the Sarakatsan nomads in northwestern Greece, norms are very strong. Campbell reports (restated in Merry, 1984, p. 283): "In one very poor family, the father failed to defend his son against the insult of being pelted with dung by another little boy, displaying to onlookers his inability to protect the honor of his family in more important ways. The same man looks dirty and uncouth despite gossip about his unkempt appearance, and his wife is criticized for laughing and joking, as it implies that she may be guilty of sexual immodesty. Although their daughter is apparently virtuous and chaste, she is tainted by her mother's immodesty, and no honorable man seeks her hand." This example raises a question about the effectiveness of norms. There are apparently normative sanctions imposed by others against various actions of this family, but the sanctions appear less effective than one might expect.

5. In Gush Emunim settlements of the West Bank in Israel, strong norms require the men of the community to form a minyan of ten or more and pray in the synagogue each day. Some men do, and some don't. Those who don't claim to be busy with their work. They stay indoors at the time when others are going to pray. This example raises the issue of the conflict between the demands of daily life and the demands of norms. How are norms maintained in the face of such conflicts? The demands of daily life are ever present and are associated with interests that have material consequences. Norms often demand, such as in cases like this, actions which are unrelated to consequences that the individual, or even the whole group, will experience. How are the norms maintained?

6. Elias (1982) traces the evolution of norms concerning table manners in French provinces, using writings about etiquette. Elias shows how these norms, which began by proscribing only the most offensive actions (such as coughing up a piece of meat at the table) evolved into the most elaborate prescriptions of what to do at the table. Furthermore, the norms differed in different social groups, and there was a hierarchy of elaboration, corresponding to social status and having its peak at the king's court. This example raises questions about how and why norms become increasingly elaborate over time, as well as why manners are more elaborate in groups of higher social status.

### Distinctions among Norms

The diversity among the examples above suggests that it is useful to make some classifications of norms. Although this cannot be fully done at this point in this exposition, a start may be made.

First, norms are directed at certain actions, which I will call focal actions. In the example about the three-year-old and the cellophane candy wrapper in Berlin, the focal action is dropping the wrapper on the sidewalk (more generally, any action that has the effect of littering the sidewalk).

Some norms discourage or proscribe a focal action, and I will call these proscriptive norms. Other norms, such as the norm to smoke marijuana among the young people at the beach house or the norm among members of Gush Emunim settlements to form a minyan and pray every day at the synagogue, encourage or prescribe a focal action. I will call these prescriptive norms. Proscriptive norms provide negative feedback in the system, damping out the focal action; prescriptive norms provide positive feedback, expanding the focal action. When there are only two possible actions, of course, one is proscribed and the other is proscribed by the same norm. For example, the norm of walking to the right when encountering another pedestrian walking in the opposite direction is simultaneously prescriptive and proscriptive. The distinction is meaningful only when the number of alternative courses of action is greater than two.

For any norm there is a certain class of actors whose actions or potential actions are the focal actions. The statement "Children should be seen and not heard" specifies a norm for which children constitute this class. I will call members of such a class targets of the norm, or target actors. There is also a class of actors who would benefit from the norm, potentially hold the norm, and are potential sanctioners of the target actors. These are actors who, if the norm has come into being, assume the right to partially control the focal action and are seen by others who would benefit from the norm to have this right. For the norm specified by the statement above, parents, or adults more generally, are those who hold the norm. It is possible that children also hold the norm, but the operation of the norm and its supporting sanctions does not depend on this. I will call those who would benefit from the norm and thus assume the right to control the target action (who are also ordinarily the potential sanctioners) beneficiaries of the norm. The current beneficiaries of the norm may be those who initiated it, or they may have merely continued the enforcement of a norm initiated by persons who preceded them.

For some norms, such as the one concerning children mentioned above, the targets of the norm and the beneficiaries are not the same persons. The norm benefits one set of actors and is directed toward actions of another set. I will label such norms disjoint norms because the set of beneficiaries and the set of targets are disjoint, resulting in a physical separation of opposing interest. The beneficiaries have an interest in the norm being observed, and the targets have an interest in the focal action being unmodified by the norm.

For many norms, however, including all those described in the earlier examples (except for the norm about not smoking marijuana held by the parents of the high-school girl), the set of beneficiaries of the norm coincides with the set of targets. In such cases, the interests favoring observance of the norm and those opposing its observance are contained within the same actors. Each actor is simultaneously beneficiary and target of the norm. I will call norms of this sort conjoint norms.

The distinction between disjoint and conjoint norms reflects only the extremes of the variations that may occur. Figure 10.2 shows those extremes, along with

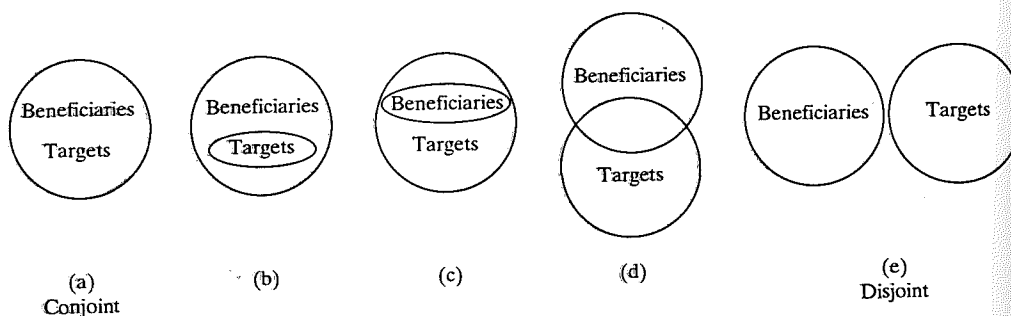


Figure 10.2 Inclusion relations of beneficiaries and targets of a norm for different types of norms.

intermediate cases. In cases b, c, and d, some persons are both beneficiaries and targets. In case b there are also some beneficiaries who are not targets. An example is the norm against sexual relations before marriage, which is held not only by unmarried persons, the targets, but also by married persons. In case c there are targets who are not beneficiaries. For example, norms about proper behavior held by members of a community are directed not only at community members as targets, but also at strangers. In case d some beneficiaries are not targets, and some targets are not beneficiaries.

A clarification of what is meant by the term "sanction" is also useful. If holding a norm is assumption of the right to partially control a focal action and recognition of other norm holders' similar right, then a sanction is the exercise of that right. A sanction may be negative, directed at inhibiting a focal action which is proscribed by a norm, or positive, directed at inducing a focal action which is proscribed by a norm. I will use the terms "sanction" and "effective sanction" interchangeably, indicating by either an action on the part of a beneficiary which has some effect in moving the focal action in the direction intended by the sanctioner.

One final distinction concerns selection of a focal action, an action to be discouraged or encouraged by a norm, from a set of mutually exclusive actions. In some cases the selection of the focal action is largely arbitrary, but in others it is not. The former is exemplified by the convention of driving on the right side of the road (or, in England and Australia, on the left). It is arbitrary whether the action defined as correct is driving on the right or driving on the left. Once the convention has been established, however, all are better off if each follows the convention. The interests in a particular direction of action depend on whether it is the action being carried out by others. If a convention has established the direction of a norm, I will call the norm a conventional norm.<sup>2</sup>

2. Ullmann-Margalit (1977, p. 97) calls these coordination norms and distinguishes between those that arise through convention and those adopted by decree. I will not make use of this distinction.

For many norms the focal action has not been arbitrarily selected. The targets' interests lie in the direction of action opposing observance of the norm, and the beneficiaries' interests lie in the direction of action favoring observance of the norm. These interests in particular directions of action would remain, whether or not the norm existed and independent of others' directions of action. In this case the direction of the norm depends on more than convention. I will call these essential norms. This last distinction can be illustrated, as Ullmann-Margalit (1977) has done and as I will do shortly, using simple payoff matrices from the theory of games.<sup>3</sup>

### The First Condition: Externalities of Actions and the Demand for a Norm

In Chapter 2 I indicated that one distinction which is important for the functioning of social systems is that between events that have consequences only for those who control them and events that have external consequences (that is, externalities) for actors who have no control over them. The latter events are intrinsically of interest to actors other than those who control them. When such events are actions, there are two kinds of externalities of actions: When an action benefits others, the action has positive externalities; when an action is harmful to others, the action has negative externalities. If an action benefits some and hurts others, then its externalities are positive for the first set of actors and negative for the second.

An action that has externalities generates interests in the action among those actors who experience the externalities. Yet there is no general way in which the consequences of the action for those affected actors can enter the utility function of the actor taking the action. Actors harmed by an action that benefits the actor in control of it experience negative externalities, as exemplified by nonsmokers sitting near a smoker. Those benefited by an action that benefits the actor controlling it experience positive externalities, as exemplified by passersby who benefit from a householder's removing snow from the sidewalk in front of his house. The problem for those other actors in the first situation is how to limit the action which is harming them (and how much to limit it). The problem in the second situation is how to encourage and increase the action (and to what level it should be encouraged).

A special case of the latter problem is that of paying the cost of a public good when each actor's action has beneficial consequences for others, by helping to bring about the public good, but the benefits to himself are less than the costs he will incur. Only if enough actors can be induced jointly to carry out the action to

3. Ullmann-Margalit distinguishes three kinds of norms, which she calls prisoner's dilemma norms, coordination norms, and norms of partiality. These correspond approximately to what I have termed essential norms, conventional norms, and disjoint norms, respectively. Essential norms, however, to use my terminology, may be disjoint or conjoint, whereas Ullmann-Margalit's three classes are mutually exclusive.

make the benefits exceed the costs for each will the public good be provided. A parallel problem exists for a public bad; for example, in overgrazing of a commons, each herd-owner's increase in grazing will increase his own benefits, but at a cost to others. Only if all the herd-owners with access to the commons can be induced to limit grazing by their animals will the grazing be reduced to the level at which the land will produce maximum nutrition.

When an action generates externalities for others, they may be able to make their interests felt through wholly individualistic means. For example, one of them may engage in an exchange with the actor whose action imposes the externalities, offering or threatening something to bring about the outcome he desires. But this may not be possible if the externalities are spread among several actors, no one of whom can profitably make such an exchange.

When exchange is possible, this gives a solution which is a special case of that described by Coase in "The Problem of Social Cost" (1960). The general solution is a market in rights of control, in which the actors who do not have control of the action may purchase rights of control from those who do, the former being limited only by their interest in the action and their resources. It is easy to see that if there are no transaction costs in such a market, the outcome will be a social optimum (which is defined only relative to the initial endowments of resources of the various parties in the market), at which no further exchanges are mutually beneficial. Those hurt by the level of action existing at the outcome would be even more hurt by parting with the resources that the actor controlling the action would take to limit it further.

In the case of a public good, each of the actors who is benefited by the actions of others would exchange rights of control of his own action for rights of partial control of the action of each of the others. For example, each resident in a town might agree to the building of a public park and to contribute an equal fraction of the cost. This constitutes a multilateral exchange in which each gives up the right not to contribute in return for the giving up of the same right by each of the others. (I put aside for the present the question of how such a multilateral exchange might be organized.)

Similar markets have been developed in regulation of environmental pollution. The amount of total pollution allowed is not set by market forces, but marketing of rights to pollute occurs among those who operate pollution-generating plants (see Noll, 1983). Yet there are many activities in society in which markets in rights of control cannot easily come into being, for one reason or another. In a social situation in which one person is smoking and another finds it irritating, the second can hardly say to the first, "How much will you take to stop smoking?" A high-school girl at a party where all others present would like her to smoke marijuana but who knows her parents would not can hardly ask for bids from the two opposed sets of others for control of her action. There is a wide range of situations in which an action has extensive externalities but a market in rights of control of the action is either impracticable or illegal.

The condition under which interests in a norm, and thus demands for a norm,

arise is that an action has similar externalities for a set of others, yet markets in rights of control of the action cannot easily be established, and no single actor can profitably engage in an exchange to gain rights of control. Such interests do not themselves constitute a norm, nor do they ensure that one will come into being. They create a basis for a norm, a *demand* for a norm on the part of those experiencing certain externalities.

The externalities created by the action may, as indicated earlier, be positive or negative. In high schools, for example, positive externalities are created by athletes who contribute to the success of a team, which in turn contributes to the school's general standing in the community (which in turn contributes to the other students' feelings of well-being or pride). Often a norm does arise, one which encourages potentially good athletes to devote their energies to interscholastic sports. In contrast, students who get especially high grades create negative externalities for other students, insofar as the teachers grade on the curve. High-performing students increase for other students the effort necessary to receive good grades, thus making matters more difficult for others. Often a norm arises in this case also; students impose a norm that restricts the amount of effort put into schoolwork.<sup>4</sup>

How a norm actually comes into being once a demand is created by externalities is a matter which I will examine in Chapter 11. But the genesis of a norm is based in externalities of an action which cannot be overcome by simple transactions that would put control of the action in the hands of those experiencing the externalities.

Several points follow from this central premise. One implication is that the potential beneficiaries of the norm will be all those who are affected in the same direction by the action. If a norm does arise, it will be those persons who will claim a right to have partial control over the action and who will exercise their claim by attempting to impose normative sanctions on the actor performing the action to induce the direction that benefits them, often at that actor's expense. A further implication is that a potential conflict of norms arises when an action has positive externalities for one set of persons and negative externalities for another. In the example of the high-school girl whose friends' approval is contingent on her smoking marijuana and whose parents' approval is contingent on her not doing so (or on their ignorance of her doing so), there are such opposing externalities. If she does not smoke, she dampens the party, destroys the consensus, and perhaps reminds some of those present of their similar normative conflicts. If she does smoke and her parents learn of it, they are made unhappy as their pride and trust in her are undercut.

The structure of interests created by externalities in which norms have their genesis may be presented more systematically by using simple situations whose

4. Of course, when academic activities are organized interscholastically, this can generate a prescriptive norm concerning studying. Striking cases of this may be found in a description of statewide competition in academic subjects among rural schools in Kentucky (Stuart, 1950, p. 90).

Table 10.1 Payoff matrix for two-person joint project.

		A <sub>2</sub>	
		Contribute	Not contribute
A <sub>1</sub>	Contribute	3, 3	-3, 6
	Not contribute	6, -3	0, 0

outcomes can be described by payoff matrices such as are used in theory of games. For example, suppose that two persons are told separately, "You may take either of two actions: contribute \$9 to a common project, or contribute nothing. For each \$3 that is contributed, an additional \$1 will be earned by the project (that is, there will be a return of \$4 for each \$3 contributed). The final total will be divided equally between the two of you, regardless of who made a contribution." Each can assess the net gains or losses for himself and for the other, for each combination of actions. These are expressed in Table 10.1, where the values of the outcome (in dollars) for each of the persons (A<sub>1</sub> and A<sub>2</sub>) are given in each cell.<sup>5</sup>

If neither contributes, there is no gain or loss for either. If A<sub>1</sub> contributes and A<sub>2</sub> does not, A<sub>1</sub>'s contribution of \$9 plus the \$3 earned will be divided equally, giving \$6 to each. For A<sub>2</sub> this will be a net gain, as indicated in the upper-right-hand cell of the table. But for A<sub>1</sub> the original \$9 contribution must be subtracted, giving him a net loss of \$3. The gain and loss are reversed for the case in which A<sub>2</sub> contributes and A<sub>1</sub> does not.

This situation creates a pair of actions, each having externalities for the other actor. As Table 10.1 indicates, A<sub>1</sub>'s action (of contributing or not contributing) makes a difference of \$6 (between 3 and -3 or between 6 and 0) to A<sub>2</sub>, and A<sub>2</sub>'s action makes a difference of \$6 to A<sub>1</sub>. Furthermore, in both cases the externalities go in the direction opposite to the actor's own interests. Each actor is better off by not contributing (whichever action the other takes), but not contributing makes the other actor worse off. Finally, the external effects of the other's action are greater for each than are the direct effects of his own action.

5. This payoff structure is that of a prisoner's dilemma. See Luce and Raiffa (1957) or Rapoport and Chammah (1965) for a discussion of this game.

A<sub>1</sub>'s action makes a difference of only \$3 to him, but A<sub>2</sub>'s action makes a difference of \$6—and similarly for A<sub>2</sub>.

The result of this situation is that each has an incentive not to contribute (since he will lose \$3 by so doing), and if both do not, each gets nothing. Yet if both did contribute, each would gain \$3. The optimal action for each actor gives a social outcome which is not an optimum. Both would be better off if both took the action which is *not* individually optimal, that is, if both contributed to the project.

Much has been written about this structure of outcomes, but most of it is not of interest here. (For references to some of this literature, see Axelrod, 1984.) What is of interest is Ullmann-Margalit's (1977) discussion of this structure as calling for or generating one type of norm, which she calls prisoner's dilemma (or PD) norms. Her argument is that such a structure of outcomes creates an incentive for all parties involved to set up a norm that will constrain the behavior of each in the direction of carrying out the action that is better for the others (in the example above, contributing to the joint project). In the terminology introduced earlier, such a structure of interdependence of actions creates externalities for each and thus an interest on the part of each in the creation of a norm.

In situations of this sort, however, where two persons' actions affect each other in the way shown in Table 10.1, a norm is not necessary at all. Either person can propose an exchange in which each gives the other rights of control to his action and gets rights of control of the other's action.<sup>6</sup> Each has resources (his own action) that are of more value to the other than the resources held by the other (the other's action). Thus by exchanging rights of control each gets something that is worth more to him than what he gives up. Each will exercise the control over the other's action in the direction which benefits himself, and in so doing will bring about a social optimum. In the example above A<sub>1</sub> will contribute A<sub>2</sub>'s \$9, A<sub>2</sub> will contribute A<sub>1</sub>'s \$9, and both will gain \$3 as the outcome.

Where there is a pair of interdependent actions for which the self-interested action of each imposes negative (or positive) externalities on the other that are greater than the benefits (or costs) that the other's own self-interested action brings, a mutually profitable exchange is always possible in principle. Logistics may, of course, preclude such exchange. In the game-theoretic analysis of the prisoner's dilemma, the possibility of exchange is excluded, because by assumption the players cannot communicate. But no such constraint is necessary here. Norms can arise only where there is communication; thus bilateral exchange is possible in all those two-actor cases where the possibility for a norm exists.

6. As far as I know, Erling Schild and Gudmund Hernes were the first (independently, in 1971) to point out that the simplest social solution to the prisoner's dilemma is exchange of control between the two players, an action which is rational for each. Bernholz (1984) has shown that Sen's paradox of a Paretian liberal (discussed in some detail in Chapter 13), where the payoff structure is that of a prisoner's dilemma, is solved in the same way. If the exchange is not instantaneous, of course, but requires a promise on the part of one or both, it becomes necessary to introduce some form of retribution.



There is an apparent exception in those cases where communication exists before and after the action, but not during the action itself. However, any agreements reached before the action or any retributions taken after the action need make no reference to a norm, but can be treated wholly within the framework of bilateral exchange—although possibly of course requiring introduction of notions of trust and mutual trust, as discussed in Chapter 5.<sup>7</sup>

The one true exception, in which the social optimum is not attained by an individualistic solution or by a bilateral exchange, is where the actions are pairwise, but the two actors are not in contact either before or after the action (or will meet only in the distant future), and thus have no opportunity either to make an agreement or to carry out the terms of a prior agreement.<sup>8</sup> In that case a norm, in which sanctions are imposed by others who are in contact with the actors after the action, can bring about a social optimum; bilateral exchange cannot.

It is best to clarify what the word "exchange" implies in the current context, for the example may otherwise be misleading. The imagery evoked by exchange in the context of this example is that one actor approaches another with an offer, "You let me make your decision, and I will let you make mine" or "Let us contribute together" or something similar. This is certainly what happens in some cases. In an examination of the emergence of norms, however, it is appropriate to conceive of a succession of comparable projects, extending over time, in which a new decision arises each time. This expands the possibilities to incorporate exchanges, implicit or explicit, that cover two or more projects (for example, "If you fail to contribute this time, I will not contribute next time"). This conception is especially relevant for those cases in which it is not logistically possible to exchange control or rights to control on a given occasion. It is also relevant for those cases in which there is no project involving simultaneous contribution but separate actions of each actor which exhibit the same pattern of internal and external effects. For example, a person must decide whether to take an action, such as watching his neighbor's house while the neighbor is away, that has a net cost for him but benefits his neighbor. His neighbor, in a similar situation, must make the same decision.

7. Whether or not such bilateral exchanges are considered to be norms is purely a matter of convention. I choose not to consider them to be norms because they do not exhibit the fundamental problem that must be overcome when no mutually profitable two-person exchange is possible but some  $n$ -person exchange is possible, where  $n$  is greater than 2. This fundamental problem, called the second-order free-rider problem, is discussed in Chapter 11.

8. This is a fundamental point on which Axelrod (1984, p. 49), who discusses the growth of cooperation between two players in iterated prisoner's dilemmas, exhibits confusion. At some points he seems to be asserting that pairwise interactions in large populations, where the same two parties will meet only very infrequently, will generate the same cooperation as found in his pairwise "tournaments." In general, however, Axelrod's work in that book demonstrates the point made here: that bilateral exchanges, explicit or implicit, are sufficient, without introduction of a norm, to arrive at a social optimum in pairwise interactions with externalities. See Coleman (1986b) for examination of social-structural conditions under which contact among individuals does not allow such agreements, implicit or explicit, to be effective.

Another type of implicit exchange, which may be more common empirically, is not precisely equivalent to those described above. If two actors have a social relationship, which as described in Chapter 12 consists of a set of obligations and expectations (assumed for the present to be symmetric), various actions by each may affect the outcome of an exchange. If  $A_1$  wants to prevent an action of  $A_2$ , which imposes a cost on him of \$6 but benefits  $A_2$  by only \$3,  $A_1$  has only to introduce into the negotiations some other event which he controls that has a cost for him of less than \$6 and a benefit for  $A_2$  of more than \$3. A promise or a threat with respect to this event may serve  $A_1$  as well as, or better than, the action which is analogous to the action of  $A_2$  he wants to control. To state it differently, one actor need not use as a sanction for another actor the same kind of action as the action he is sanctioning. For example, if one actor is late for a meeting, the other need not show up late for the next meeting; he can express disapproval, or he can threaten to break off the meetings altogether (if the meetings are of sufficient interest to the first actor to make this a credible threat). The other events may include some for which the costs to the sanctioning actor are very small yet the other actor's interest may be sufficiently great that the sanction is effective.

It is important to recognize these additional possibilities that actors may have for sanctioning one another because they lend importance to the existence of other events linking the actors. Attention to these additional possible sanctions is also important because of the potential asymmetries in sanctioning that may result from inequalities in actors' control of events of interest to others.<sup>9</sup>

### *Systems of More Than Two Actors*

It is when pairwise exchanges cannot bring about a social optimum that interests in a norm arise. This may be illustrated by expanding the joint project described earlier to a common project involving three actors. Again, each has the alternative of contributing \$9 or nothing. For every \$3 contributed, the product will be \$4. The total product will be divided equally among the three. Table 10.2 shows the outcomes for each combination of actions. Since the situation is symmetric for the three actors, these outcomes can be summarized more compactly, as shown in Table 10.3.

The situation here is fundamentally different from that shown in Table 10.1. It is not possible for two of the three actors to exchange control over their actions and gain by so doing. If there are no contributions, giving no net gain or loss to each, and then  $A_1$  exchanges control with  $A_2$ , each contributing for the other,

9. Another source of asymmetry is hidden by the symmetry of this example. Even for activities for which all actors' similar actions impose externalities on the others, the externalities may be unequal, providing sanctioning opportunities for some actors that do not exist for others. This is related to interpersonal comparison of utilities, and, as will be evident later, a correct untangling of that issue will be important for the analysis of norms as well as other aspects of the social system.

Table 10.2 Payoff matrix for three-person common project.

		A <sub>3</sub>			
		Contribute	Not contribute		
A <sub>1</sub>	Contribute	A <sub>2</sub> Contribute	A <sub>2</sub> Not contribute	Contribute	A <sub>2</sub> Not contribute
	Not contribute	Contribute	Not contribute	Contribute	Not contribute
	Contribute	3, 3, 3	-1, 8, -1	-1, -1, 8	-5, 4, 4
	Not contribute	8, -1, -1	4, 4, -5	4, -5, 4	0, 0, 0

they end up losing \$1 while A<sub>3</sub> gains \$8. If A<sub>3</sub> is contributing, A<sub>1</sub> and A<sub>2</sub> each gain \$4 without an exchange. If they exchange control, with each contributing for the other, the gain for each is \$3, making each of them \$1 worse off than they would be without the exchange.

Only if both A<sub>2</sub> and A<sub>3</sub> can be induced to change their actions from not contributing to contributing, contingent on A<sub>1</sub>'s contribution, does it become profitable for A<sub>1</sub> to join in such an arrangement. In such a case the outcome for each changes from no gain to a gain of \$3. Thus a compact among the three is necessary to bring about a gain to each. One form of compact is a norm, by which the right to contribute or not is no longer held by each actor, but for each is held by the other two. It is in this way that it can be said that each comes to have interests in a norm.

The structure of interdependence in this case is one in which, if a norm arises at all, it will be a conjoint norm, with the same actors being targets and beneficiaries. It will be an essential norm, not a conventional one, because there

Table 10.3 Summary of outcomes for three-person common project.

Number of contributions	Gains or losses to—	
	Contributors	Noncontributors
0	—	0
1	-5	4
2	-1	8
3	3	—

is one direction of action that benefits each (contributing) and one that does not. It would be possible to construct a similar artificial example and matrix of outcomes for which interdependence would generate interest in a conventional norm. But that is straightforward and self-evident, and I will not present it here.<sup>10</sup> For a disjoint norm the matter is somewhat different, and I will put aside examination of such norms until a later point. It is, however, useful to examine a question that arises concerning some conventional norms, where the externalities imposed by one actor's action on the others are not immediately apparent.

*Do Norms Arise Only When There Are Externalities of Actions?*

There are some norms which seem not to be generated by an action's imposition of externalities on others. For example, in high schools certain subgroups of girls or boys will have strong norms about how their members dress.<sup>11</sup> In the 1950s ducktail haircuts constituted observance of a norm by certain groups of boys. Certain groups of girls wore bobby socks (or even a particular color of bobby socks); certain groups of boys wore white bucks, and others black leather jackets. In Jerusalem some women keep their heads and arms covered, and some men wear black yarmulkes, reflecting membership in an orthodox Jewish community characterized by a particular set of religious observances. In Cairo some women dress wholly in black with heavy black veils covering their faces, reflecting membership in a Muslim community characterized by a particular set of religious observances. In rural Pennsylvania some women wear bonnets and plain-colored clothes without buttons, reflecting membership in an Amish community characterized by a particular set of religious observances.

10. A table of payoffs that can generate a conventional norm, exemplified by the norm of walking to the right or to the left when passing on a sidewalk, is as follows:

		A <sub>3</sub>			
		Left	Right		
A <sub>1</sub>	Left	A <sub>2</sub> Left	A <sub>2</sub> Right	Left	A <sub>2</sub> Right
	Right	Left	Right	Left	Right
	Left	0, 0, 0	-4, -6, -4	-4, -4, 6	-6, -4, -4
	Right	-6, -4, -4	0, 0, 0	-4, -6, -4	0, 0, 0

11. See *The Adolescent Society* (Coleman, 1961) for various examples among high-school students.

All these norms about dress are conventional norms, despite the doctrinal rationale for the direction that some of them take. But how is it that a member's action in conformity with the particular norm creates a positive externality for other members of the group? For persons who have come to constitute a group and want to differentiate themselves from others, common dress constitutes a very efficient means of doing so. Each member's obeying the norm strengthens the expression of group solidarity and the differentiation from others. Observance of rules about dress is similar to observance of dietary norms, rules of etiquette, and other differentiating characteristics (see Goode, 1960; 1978).<sup>12</sup> Observance by fellow members aids and supports each member, and failure to observe constitutes a threat to the solidarity of the group. This is an instance in which a derivation from theory can aid research. Measurement of the strength of dress codes and their degree of observance in a particular subgroup can show the strength of interest of the members in subgroup membership.

Dress codes illustrate a form of conventional norm for which externalities do not exist prior to and independent of the norm. They contrast with conventions such as driving on the right, for which negative externalities exist in the absence of the norm and the norm provides benefits by reducing them. Dress codes exist where the strength of members' interest in group membership is sufficiently great that an opportunity for positive externalities exists. The dress code makes possible those positive externalities by prescribing dress that will declare one's group identity to other members and to nonmembers.

#### *Status Groups, Norms of Etiquette, and Standards of Speech*

Norms of etiquette, such as those studied by Elias (1982) and mentioned earlier in this chapter, are somewhat different from norms about dress maintained by a group such as a religious group. As Elias showed, elementary norms of etiquette are essential norms, constraining the target actor's behavior so that it is attentive to the interests of those interacting with that actor. At the same time, however, norms of etiquette create a status group composed of those who conform to them. Because their actions attend to the interests of others in the vicinity, they can make a claim to be "better" than those who do not observe the norms. Action in conformity with a norm of etiquette creates a positive externality for members of the status group, who hold the norm, by differentiating them from those who do not hold it—just as in the case of religiously prescribed dress codes. Since membership in the status group is defined by conformity to the norm, however, anyone who acts in conformity with the norm can enter the group. The norm will not give a positive externality for members of the group unless conformity to the norm is sufficiently difficult that outsiders cannot easily enter the group.

12. Dietary restrictions may be more than conventional. Some have arisen for reasons of health or reasons of scarcity. For an examination of the use of rules of etiquette in differentiating one group from another, see Elias, 1982.

Thus the norm, which at the outset only constrains actions having negative externalities, has the potential to induce actions that bring positive externalities by creating a status group of those who hold the norm. That potential is realized, however, only if the norm is elaborated in such a way to make entry into the status group difficult. If a set of actors is capable of establishing a norm of etiquette to meet the demand for reducing negative externalities in interaction, then this set of actors, comprising a status group, is also capable of elaborating the norm to maintain the distinctiveness of that group. (The question of whether the set of actors will be able to establish a norm of etiquette is, of course, an open one, which cannot be answered merely by specifying that persons have common interests in eliminating negative externalities or encouraging positive ones. I will examine that question in Chapter 11.)

It is not only norms of etiquette that can be generated and used by status groups as described above. Standards of speech, elaborate norms of dress, or norms of fashion are used in the same way. A social system may contain a hierarchy of status groups, in which members of a group that is neither at the top nor at the bottom attempt both to conform to the norms of the next higher group and to maintain the norms of their group in order to keep out those below.

What I have outlined in this section indicates that sets of persons develop norms not only to serve as protective devices against actions that impose negative externalities, but also to perform positive services. This is more conjectural than much of the theory in this book and clearly requires empirical study, in order to test the theory and elaborate its details.

#### *A Note on the Concept of Function in Social Theory*

In the preceding examination of the use of norms by sets of actors, I have avoided using the term "function," although it would be natural to have written, for example, that "sets of persons develop norms not only to serve a protective function against actions that impose negative externalities, but also to perform positive functions for them." I have avoided using the term because of the confusion surrounding its use in social theory. In particular, radical versions of functional analysis have purported to explain the existence of a phenomenon by its function. In this context that would mean explaining the emergence of a norm by the functions it serves for the set of actors who hold it.

It should be clear, however, that the functions a norm serves for those who hold it, or, in the terms I have used, their interests in the norm, are not sufficient as an explanation of its emergence or continued existence. The fact that a set of actors is interested in gaining the right to control the actions of individual actors is not sufficient as an explanation for their coming to gain that right. In the explanation of the emergence of norms given in this book, that is only the first of two necessary conditions. The condition under which those interests will be realized, to be examined in the next chapter, is the second half of the explanation.

The common tendency of many theorists carrying out "functional analyses" to explain a phenomenon solely by its function is the principal failing of functional analysis as a theoretical paradigm. For a theorist to go beyond this, to examine how a phenomenon has come into existence, requires going from the macrosocial level down to the level of actors, thus abandoning the paradigm of functional analysis for a paradigm that, like the one used in this book, contains actors and a theory of action.<sup>13</sup>

### What Constitutes Social Efficiency?

Most of the analysis of the preceding sections has been concerned with conjoint norms, for which the set of persons experiencing externalities from the focal action is the same as the set of persons who carry out the focal action and thereby impose externalities on others. In such a circumstance and with a group which is homogeneous, exemplified in Table 10.2, the concept of social efficiency, or a social optimum, is straightforward, because interests for and against the action are contained within the same actors. If the benefits that each obtains from his own action are less than the costs imposed by all others' similar action, then the emergence of a norm is socially efficient. All will be better off if there is a norm.<sup>14</sup> The achievement of social efficiency in this case lies in a redistribution of rights of control over each action to the group as whole.<sup>15</sup> But with a group that is not homogeneous, it becomes more problematic to define social efficiency. The balancing of positive and negative interests in each actor's action is no longer straightforward. The difficulty is most easily seen by turning to disjoint norms, where the externalities are imposed by a set of actors disjoint from those experiencing the externalities, who are the potential holders of the norm.

How can one say whether a norm is socially efficient? In the case of a conjoint norm within a homogeneous group, the establishment of the norm either brings a gain to each member and is socially efficient or brings a loss to each and is inefficient. But establishment of a disjoint norm makes the beneficiaries better off (through the rights they gain) and the targets worse off (through the rights they lose). At this point Coase's 1960 paper on the problem of social cost becomes directly relevant. Coase was not concerned with norms but with the question of how cases involving externalities imposed by one actor on another should be decided in law. Who should pay the social costs? That problem is

13. See Nagel (1970) and Stinchcombe (1968) for examinations of the logic of functional analysis.

14. By a socially efficient state or a social optimum, I mean the analogue to economists' notion of economic efficiency. I do not mean the much weaker concept of a Pareto optimum. For a disjoint norm, both the existence of the norm and its absence are Pareto-optimal points, since no move from either state to the other can be made without hurting either the beneficiaries of the norm or its targets. Only one of the two states, however, is socially efficient. See Chapter 2 and Chapter 29 for fuller discussions of social efficiency.

15. This statement is given greater precision in Chapter 30.

closer to the problem of disjoint norms than is at first apparent—for law and social norms are close relatives in the family of social control mechanisms, and the structure of the situation examined by Coase (one actor or set of actors imposing externalities on a different set) is, according to the theory presented in this chapter, exactly that in which a demand for a disjoint norm arises.

Coase (1960) argued that legal allocation of rights to the actor or actors carrying out the activity that produced the externalities or to the actor or actors experiencing them was irrelevant to the question of whether the activity would continue (the Coase theorem). If the activity was economically efficient, it would continue (with rights to impose the externalities purchased by the actor producing them, if necessary); if the activity was not economically efficient, it would not continue (with rights to produce the externalities purchased from the actor producing them, if necessary). There are conditions attached to the Coase theorem, such as absence of transaction costs, the existence of a market value for the costs imposed by the activity, and negligible effects of the legal allocation of rights on the distribution of resources within which economic efficiency must be defined. The theorem is directly relevant to the case of disjoint norms, however, for Coase's point is that there is an economic value to the activity that produces the externalities and an economic cost created by the externalities, and if (but only if) the former exceeds the latter, the activity will continue, independent of who pays the costs. The possibility of transactions outside the context of the externalities themselves (that is, the possibility of purchase of rights by one party or the other) is what brings this about.

The implication of this theorem for disjoint norms is that if the externalities imposed by the activity are sufficiently great that the persons experiencing them are motivated to induce the actor to cease, and if those persons have the resources to make that motivation effective, then they will do so, whether or not a norm is in existence (that is, whether or not the informal rights to control the action are held by them). This implies that a norm (or a law) is superfluous and would make the existence of norms or laws inexplicable.

How is the logic of Coase's argument consistent with the existence of norms and laws? The answer lies in the assumption on which the Coase theorem is based—that there are no transaction costs in inducing the actor to cease carrying out the action when he has the rights.<sup>16</sup> There are, however, transaction costs in real social systems, and a norm can reduce the transactions necessary to achieve socially efficient outcomes. When rights are held by those experiencing the

16. There are other problems with the Coase theorem. If rights allocation does not matter, then the right to steal or to kill without fear of sanction should lead to outcomes no different from the outcomes arising in the absence of such rights: The potential victim need only buy off the predator at a mutually satisfactory price. In such a social system physical strength and access to weapons become the principal resources of value. More generally, the Coase theorem neglects the fact that rights allocations can affect the incentive to impose externalities on others (for example, in the form of threats); it treats externalities as necessary by-products of activities carried out for other purposes.

externalities and those rights are recognized by the target actor, then the action is inhibited without any transaction occurring. Sanctions are unnecessary and will be applied in only two circumstances: when a target actor misinterprets the situation and incorrectly believes his power is sufficient to allow him successfully to ignore the norm; and when the target actor does have sufficient power, but the norm holders misinterpret the situation, believing they can successfully sanction the action.

### *The Importance of Other Resources for Disjoint Norms*

Enforcement of what is socially optimal through the reallocation of rights and the threat of sanctions implies the possibility that the actors who experience the externalities can use their control over events *other than* the actions creating the externalities—events in which the target actor has an interest. It is their control of such events that allows the events to be used as sanctions or potential sanctions, and it is the importance (or “value”) of those events that determines whether the sanctions will be effective.

When the social optimum is that an action not take place, the action will be constrained and the social optimum achieved only if those experiencing the externalities have the requisite social relationships to both bring a potentially effective norm into existence and make it effective. I will discuss this problem in the next chapter and will show some of the ways in which the potentially effective demand for a norm may be realized. The absence of social relationships among those experiencing the externalities will prevent the social optimum from being attained, if that optimum is observance of the relevant norm.

There is, however, a more important qualification, which concerns the existing distribution of rights and resources, that is, of control of events by actors. Just as in the case of economic efficiency, the social optimum is defined relative to an existing distribution of rights and resources (see Chapter 30). If that distribution is highly unequal, this implies that the interests of some actors count for much more than do the interests of others. Those actors having more power will be able to impose disjoint norms to govern the actions of those having less power, and they will be able to counter effectively norms imposed to govern their actions and to resist sanctions applied to enforce those norms.

This situation is illustrated by two kinds of examples from traditional societies. One concerns norms governing the actions of women. In many such societies there are stringent and effective norms governing the behavior of unmarried women, the behavior of married women, and the behavior of widows, but not similar norms governing the behavior of men. The target actors are women, but the actors benefiting from the constraints and ensuring that sanctions will be applied to violators are generally men of all ages and women older than the target actors, that is, actors other than the targets. Do these norms, together with their observance, give a social optimum? Social anthropologists with a functionalist

orientation argue that they do (see Gluckman, 1955; 1963). That answer is correct, but only for the distribution of power that exists in such societies, a distribution in which men and older women have a great deal of power. If that highly unequal distribution of power were not taken as given, it could not be said that the norms result in a social optimum. (See Chapter 30 for further discussion.)

The second example that illustrates this point is the fact that even when a norm exists for which they are among the target actors, wealthy and powerful members of traditional societies are often not subject to effective sanctions. Their actions are less impeded by the norm because, given the distribution of power, the sanctioner or sanctioners cannot apply the sanctions without losing more than is gained by bringing about observance of the norm. This result is not due to an inability to share the cost of the sanction; it occurs because the total costs of applying the sanction to a powerful person are greater than the benefits achieved by the sanction, merely because of the differential power.

### *Do Similar Results Hold for Conjoint Norms?*

In the case of disjoint norms, control over events other than the focal action is important. What about conjoint norms? In examining those norms, using the example of the common project, I directed attention principally to the focal action itself, with the social optimum defined only in terms of those actions which produced a public good or public bad. The implications of the preceding section, however, are that *all* the events controlled by target actors or by beneficiaries of the norm that are of interest to actors in the system become relevant in defining the social optimum. Analyzing conjoint norms as if the only resource an actor has at his disposal is the action of contributing or failing to contribute to a common project ignores those other events.

The resources relevant to defining the effective demand for a norm consist of actions that have nothing to do with the norm. They are primarily disapproving statements and withdrawal of respect, actions which are of differential interest to a target actor depending on who the sanctioner is and who the target actor is.

Consider an example introduced earlier, the one about tea bags. Assume that observance of an antihoarding norm makes each person better off.<sup>17</sup> The norm is a conjoint norm. Suppose that there is one person in the organization who has sufficient power (the boss) that his disapproval of a sanction would outweigh the benefits of the sanction; there is no distribution of sanctioning costs that will make the benefits for each outweigh the costs to each. If that is the case, this actor, the boss, could hoard tea bags without being sanctioned, but all other tea drinkers in the organization would be sanctioned by one another. Only if the

17. This may not be the case. Hoarding merely transfers the common stock to individual stocks. But absence of the norm may reduce availability in another way, since individuals may use stock from their hoards elsewhere.

boss had internalized a norm against such an action would he not hoard. This scenario bears a striking resemblance to those traditional societies in which the wealthy and powerful are able to ignore certain norms without being sanctioned.

The question, then, is this: Is a social optimum achieved if the norm is violated by this one actor, through an unsanctioned action which reduces availability of tea bags for all others? The answer is yes, just as it was for the case of a disjoint norm discussed earlier. The social optimum is the state in which the powerful actor does not ever experience reduced availability of tea bags, but even has them during temporary shortages, and each of the others experiences slightly reduced availability. But, just as in the example from traditional societies, this social optimum is based on the existing distribution of rights and resources (that is, control over events) held by members of the organization. The statement that a social optimum is achieved when the availability of tea bags to the average actor is maximized is incorrect if it contains the implicit assumption that the availability of tea bags to each actor should be weighted equally in arriving at the availability to the average actor. Only if the powerful actor is given a weight proportional to his power in the system is there a correct definition of availability. In the mathematical model described in Part V, the differential weights are derived from the power of different actors: the equilibrium (that is, the socially efficient state, or the social optimum) is based on the initial distribution of resources.

Certain implications that are not obvious follow from the differential power of actors to exercise sanctions and to violate a norm. Perhaps the most straightforward is the implication that those who because of greater power would be likeliest to impose a sanction on others (because an equally effective sanction would cost them less) would also be those most likely to violate the norm (because others would have less power to sanction them). The same implication follows from differential interest in the focal action. For example, a person for whom the availability of tea bags is especially important should be more likely to violate a norm against hoarding and should also be more likely to sanction others for violating such a norm. There is, however, the fact that violating a norm can put one at a disadvantage with respect to sanctioning others, so norm violation and sanctioning tend to preclude one another. Thus it is probably more correct to state these implications as follows: Norm violation and sanctioning should both be positively associated with power in the system, although possibly not with each other; and norm violation and sanctioning should both be positively associated with interest in the focal action, although possibly not with each other.

The various issues surrounding the emergence and observance of norms and sanctioning as discussed in this chapter are difficult to describe precisely with words. This is especially true for issues involving relative power and the use of resources that go beyond the focal action. The results given here, together with others that require formal derivation, are shown more precisely in Chapter 30, where these questions are reexamined with the aid of a formal model.

## Systems of Norms

Throughout this chapter I have examined norms one by one, as though they existed independently. Some norms are relatively independent, but this is not generally true. Norms, like laws, are related to one another; in some cases these relations are hierarchical and in some cases due to overlapping jurisdictions. The exploration of such relations constitutes a major task in itself, one that is beyond the scope of a treatment of the foundations of social systems.

A question that necessarily arises in such an investigation is whether an analysis of systemic relations of norms must make the descent from macro to micro level and the ascent from micro to macro level. That is, must actors be brought in as intervening elements, or can a system of norms be analyzed as relations among norms per se? The question is comparable to one concerning the relations among laws that make up a system of common law: Does the study of common law have to go outside the body of laws itself and examine the actions of individual actors? One answer to this question is obtained by observing the way construction of legal theory is actually carried out. It is carried out by examination of cases, cases involving actors taking actions, other actors experiencing the consequences of those actions, and judges giving justifications for the resolution of disputes about liability for those consequences. In carrying out such examination, legal theorists infer a general principle (that is, the common law) from the resolution of specific cases. Each of these cases involves a descent to the micro level, even though the law itself, the principle as it has evolved, is a system-level property.

In the examination of jurisdictional questions and questions involving conflicts between different laws, legal theorists proceed in the same way, carrying out examinations at the micro level of specific cases of actors and events, control and interests, in order to make generalizations about the relations between laws.

These purely methodological issues do not enter into the substance of theory concerning systems of laws or systems of norms. The conclusion I draw from the character of legal scholarship is that research on systemic relations among norms is also best carried out not through analyses (whether formal or not) of the abstract content of norms, but by moving back and forth between micro and macro levels at each interface between two norms, examining the way the conflict (or other contact) between norms comes about and how actors resolve normative conflict.

« 11 »

## The Realization of Effective Norms

In the preceding chapter I examined the conditions which lead to the demand for a norm, that is, to interest in the creation of a norm and in the imposition of sanctions to bring about its observance. I said nothing about the conditions which allow this demand to be realized by bringing into being a norm and sanctions. The question which must still be answered is this: What is required to get from interests in a norm to the actual existence of a norm backed by sanctions?

It is useful to clarify what I mean by the qualification "backed by sanctions." In the preceding chapter I defined the existence of a norm as the state in which a socially defined right to control an actor's action is held not by the actor but by others. If a norm exists, it may be generally observed by target actors even though they find it against their immediate interests to do so. Sanctions may seldom be necessary. But unless the beneficiaries of a norm have the capability of applying effective sanctions when necessary, it is meaningless to say that they hold the right to control the action. (To say that there is an effective sanction does not imply that the sanction is always effective or effective for all target actors, but that it is effective for at least some target actors some of the time.) Thus, when I use the terms "effective norm" and "effective sanctions," I mean that the potential for enforcement exists for at least some of the focal actions.

The fundamental problem exhibited in the common project involving three actors used as an example in the preceding chapter is one of social organization. In the two-actor project (Table 10.1), each person has the resources to prevent the other from imposing negative externalities on him (or, equivalently in this case, to induce the other to act in a way that brings about positive externalities).<sup>1</sup> This is not so in the three-actor project (Table 10.2). No single actor can exchange control with a single other to their mutual benefit. The externalities of the actions of each for any one of the others are less than each actor's own effect on his gains. If a social optimum is to be achieved in such a case, something beyond pairwise exchange is necessary.

For a project like that of Table 10.2, one solution is a sequence of pairwise

1. As indicated earlier, when there are only two alternative actions, as in this case, there is no distinction between prescriptive and proscriptive norms.

exchanges among the three actors ( $A_1$ ,  $A_2$ , and  $A_3$ ), in which actors  $A_1$  and  $A_2$  first exchanged rights of control, and then  $A_2$ , having the right to control  $A_1$ 's action, exchanged *this* for the right to control  $A_3$ 's action. After these exchanges the rights of control are distributed as follows (where  $E_i$  refers to the contribution of  $A_i$ ):

$A_3$  controls  $E_1$   
 $A_2$  controls  $E_3$   
 $A_1$  controls  $E_2$

If these exchanges took place, then each actor would exercise the control he possessed in a way that benefited him (as well as one of the other two):  $A_3$  would commit  $A_1$ ;  $A_2$  would commit  $A_3$ ; and  $A_1$  would commit  $A_2$ . But the first exchange, between  $A_1$  and  $A_2$ , would take place only if both actors knew that a second exchange was possible—for without that further exchange each would be giving up something worth more to him than what he received. Furthermore, after the exchange between  $A_1$  and  $A_2$  had been made,  $A_3$  would find it *not* to his benefit to exchange control with either. Thus the transactions would end after the exchange between  $A_1$  and  $A_2$ , and both would end up losing while  $A_3$  gained.

This solution depends not only on  $A_3$ 's willingness to carry out the second exchange, but also on a condition often not found: the knowledge on the part of  $A_1$  and  $A_2$  that further transactions will be available to make an initially unprofitable exchange a profitable one. As is evident from the study of primitive systems of economic exchange, the development of such a sequence of exchanges (by which objects come to have a value in exchange apart from their utility for the actor, leading the actor to acquire them for further exchange) is not simple (see Einzig, 1966).

### An Action-Rights Bank

The similarity of the fundamental problem of the three-actor project to one confronted in economic markets where there is no medium of exchange or central bank raises the question of whether there are other possible solutions analogous to what has come about in economic markets. One possibility is suggested by the fact that the rights of control of each actor's actions have for him a particular value, as well as having a particular value for others in the system. For example, in the project of Table 10.2 the contribution of  $A_1$  has for him a value of (minus) \$5, and it has for the other two actors, taken together, a value of \$8. Suppose all actors agree that each actor will deposit in a central "bank" the right to control his action in this event. Then this action-rights bank issues two types of shares in the right to determine  $A_1$ 's action, which it sells at \$1 per share: shares for the right to use the right in a positive direction, that is, to make a contribution; and shares for the right to use the right in a negative direction, that is, not to make a contribution. Each actor in the system may offer to buy any number of shares in each right. Whichever set of shares, positive or

negative, is more heavily subscribed is sold to the offerers, who pay the central bank (the actors on the other side get their money refunded) and receive the positive or negative right.

In the specific case being considered, the right to control  $A_1$ 's action in a positive direction is worth \$4 to  $A_2$  and \$4 to  $A_3$ . Assume that  $A_2$  and  $A_3$  make an agreement that each will offer to buy \$3 worth of shares. That \$6 is more than it is worth to  $A_1$  to offer for negative shares; they are worth only \$5 to him. Thus if he has bought \$5 worth of negative shares, his \$5 is refunded to him, the positive right is given to  $A_2$  or  $A_3$  (it doesn't matter which since it must only be used to bring about a positive action), and the \$6 is given to  $A_1$ . Next  $A_1$  is required to contribute \$9, as a result of the positive right's being exercised by  $A_2$  or  $A_3$ . Then each actor gets \$4 back as a return from that contribution. The result is that  $A_2$  and  $A_3$  gain \$1 each, having paid in \$3 and recovered \$4, and  $A_1$  gains \$1 (\$6 - \$9 + \$4). Similar transactions involving rights to control  $A_2$ 's and  $A_3$ 's contributing would result in  $A_1$ ,  $A_2$ , and  $A_3$  each gaining \$3 overall.

Such a system would create a social optimum through a device comparable to a central bank and a market with a medium of exchange. The procedures of the action-rights bank have, in effect, overcome the public-good problem confronted by  $A_1$ ,  $A_2$ , and  $A_3$ . The mental construction of such a device gives some insight into the character of the problem to which norms are addressed.

Although such an action-rights bank may seem strange, a majority voting system in which actors vote for passage or defeat of a project has striking similarities to it. In both cases each actor has partial rights of control over the action at issue. The depositing of rights of control over an action in the action-rights bank is analogous to the constitutional decision that certain rights will be held collectively rather than individually. These rights of control are then partitioned into votes, or shares of partial control. In casting a vote, a voter is exercising a right in the direction of a positive outcome or a negative outcome, as in the case of the action-rights bank.

There are, of course, also differences. In collectivities where actions are decided on by majority vote, shares in the right to control the action are distributed on a basis whereby each qualified member or citizen receives one share, not through the sale of shares. A member or citizen cannot alter his fraction of control over the event, as an actor can in the case of the action-rights bank, except by casting or not casting his vote, an action which has trivial costs for him. This has two implications: His control over the event is independent of his resources, that is, his wealth; and his control over the event is independent of his interests in it. (In certain cases the costs in time are not trivial. This makes voting dependent both on interests [positively] and on earnings [negatively]. For an examination of such a case in India, see Goel, 1975.) This suggests possible directions of modification of voting systems in ways that would make each voter's control of the event dependent on his interests but still independent of his wealth. (Tideman and Tullock, 1976, and Groves and Ledyard, 1977, have proposed a "demand-revealing" method for arriving at a social choice that has the

first of these properties. See Margolis, 1982, for criticism of their device.) All of this, however, goes beyond the subject of norm emergence being addressed here.

There is an alternative sometimes used by individuals who anticipate benefits from a common activity but have difficulty in overcoming the free-rider problem. They may vest rights of control over their actions in a leader, creating a conjoint authority system. This requires, of course, a high level of trust in the leader to act in terms of the followers' interests, trust which sometimes is placed when a potential leader is viewed as having charismatic qualities.

### Social Relationships in Support of Sanctions

In the absence of an externally imposed solution to the public-good problem, some kind of combined action is necessary if a social optimum is to be attained. The combined action can be the mutual transfer of rights that constitutes establishment of a norm; but for the norm to be effective there must also be an effective sanction to enforce it, if any of the actors should give indications that he will not contribute. This in turn depends on the existence of a social relationship between two actors affected by the actions of a third. Fig 11.1 shows two cases: In part a actor  $A_1$ 's action has an effect on  $A_2$  and  $A_3$  (as shown by the arrows), who have no social relationship with one another. Their social relations are with other actors,  $A_4$  and  $A_5$ . In part b there are the same effects of  $A_1$ 's actions, but actors  $A_2$  and  $A_3$  have a social relationship (the content of which I will discuss shortly).

In the case depicted in Figure 11.1(a), any sanction by  $A_2$  or  $A_3$  to direct  $A_1$ 's action so that it is not inimical to their interests must be applied by either independently. As is shown for the three-actor common project in Table 10.2, neither can do so: A threat to  $A_1$  by  $A_2$  not to contribute if  $A_1$  fails to contribute hurts  $A_1$  by only \$4, and  $A_1$ 's not contributing gains him \$5. In contrast, as shown in Figure 11.1(b), a social relationship between  $A_2$  and  $A_3$  may make it

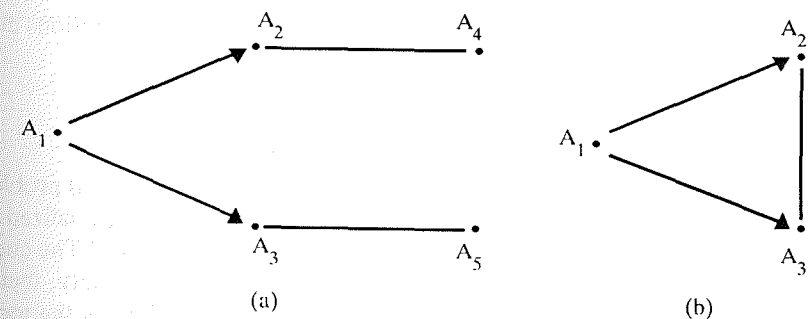


Figure 11.1 Structures of relations among actors that have differing potentials for the emergence of a norm.



possible to impose a sanction on  $A_1$  through some form of joint action that neither  $A_2$  nor  $A_3$  could impose alone. Similarly, if there exists some social relationship between  $A_1$  and  $A_2$ , it may be possible to impose a sanction on  $A_3$ , and similarly with a relationship between  $A_1$  and  $A_3$ .

But two related questions arise: First, what kind of sanction might be applied which could not be applied by either of the actors separately? Second, what is meant by the unanalyzed term "social relationship"?

There are two aspects of a social relationship that can lead to an effective sanction. One is simply communication that allows the possibility of joint action. If  $A_2$  and  $A_3$  pool their contributions, they together offer a sanction that is effective toward  $A_1$ : Their combined contribution makes a difference of \$8 to  $A_1$ , and his own failure to contribute benefits him by only \$5. Thus a joint threat by  $A_2$  and  $A_3$  to not contribute is sufficient to bring about a contribution from  $A_1$ .

Second, the social relationship between  $A_2$  and  $A_3$  may contain some other possibilities, some interests and control which give one or both actors leverage over the other. These are the obligations and expectations examined in Chapter 8 as components of a certain kind of system of trust. Because social relationships consist of obligations and expectations, held either asymmetrically or symmetrically, and because each actor continues to control some events in which the other is interested, there exists inherently in each social relationship leverage which can be used for the purpose of developing sanctions. If, in the case depicted in Figure 11.1(b),  $A_2$  has obligations toward  $A_3$ , then  $A_2$  may pay off a portion of those obligations by sanctioning  $A_1$ . But he can do this only if he can threaten  $A_1$  with some event that could make a difference of \$5 or more to  $A_1$ .  $A_2$ 's own contribution makes a difference of only \$4, which is not enough.  $A_2$  must have some obligation toward  $A_1$  that he can threaten not to honor, if his sanction is to be effective. Even a threat which costs him more than he benefits from  $A_1$ 's contribution may be viable, because of the compensation he receives from  $A_3$ . Or if  $A_2$  has control of some event in which  $A_3$  is interested (which may be nothing more than  $A_2$ 's approval),  $A_2$  and  $A_3$  can carry out an implicit exchange, in which  $A_3$  incurs the cost of sanctioning  $A_1$  in return for control of the event currently controlled by  $A_2$ .

This use of social relationships to facilitate the employment of sanctions constitutes a solution to a general problem, to which I now turn.

### *The Second-Order Public-Good Problem for Norms*

The sanctioning problem has been called the second-order public-good problem or the second-order free-rider problem. The problem can be conceptualized by considering one of Aesop's fables, known as "The Mice in Council." The council meeting was called to discuss a problem faced by the mouse society, that of how to control the cat who was slowly decimating the population. In the terms used in this book, the cat's action was imposing severe externalities on the mice and constituted, in effect, a public bad, creating constant danger for each mouse. This is the first-order public-good (or in this case public-bad) problem.

The second-order public-good problem is indicated in the statement of the wise old mouse who finally rose in the council, after a proposed solution (that a bell be put around the cat's neck to warn of its approach) had been roundly applauded. He suggested that the council consider how the bell was to be fastened about the cat's neck and who would undertake the task. The second-order public-good problem lies in the fact that, just as the cat's action imposes externalities on all, an effective sanctioning of the cat's actions also has externalities (positive in this case) for all those experiencing the benefits of the sanction; yet the benefits to the mouse who would undertake to bell the cat would not be sufficient to overcome the costs.

For the case of the three-actor project shown in Table 10.2, the first-order public-good problem lies in the fact that each will benefit only from the contribution of others; and the second-order public-good problem is that if  $A_1$  does not contribute, then the sanctioning of  $A_1$  is a public good for  $A_2$  and  $A_3$ , but neither receives sufficient benefits from his own sanctioning action to compensate the costs of sanctioning  $A_1$ . The problem may not appear to be a serious one for the three-actor common project. The second-order public-good problem for sanctioning one actor's failure to contribute to a three-actor common project is reduced to a two-actor joint project. This may be solved whenever there is the possibility of exchange between the two actors who experience externalities from the third. One can compensate the other for the net costs of applying the sanction (in this case, for instance, the costs to  $A_2$  of sanctioning  $A_1$ , less the benefits that  $A_2$  will derive directly from the effects of the sanction). More generally, the second-order public-good problem of sanctioning always involves one actor less than the first-order public-good problem.

The sanctioning problem for the case in Table 10.2 is shown in Table 11.1, where it is assumed that  $A_1$  proposes not to contribute and that  $A_2$  and  $A_3$  have contributed. Since  $A_1$  gains \$5 by not contributing (the difference between \$3

Table 11.1 Payoff matrix for two potential sanctioners in three-person common project.

		$A_3$	
		Sanction	Not sanction
$A_2$	Sanction	.5, .5	-2, 3
	Not sanction	3, -2	-1, -1

and \$8), it will cost  $A_2$  or  $A_3$  (or both together) whatever is equivalent to \$5 for  $A_1$ . Making up the \$5 for  $A_1$  is not possible for either  $A_2$  or  $A_3$  to do within the framework of the common project, since either alone can only make a difference of \$4 to  $A_1$  by contributing or not contributing. It is further assumed for Table 11.1 therefore that  $A_2$  and  $A_3$  each have a relation to  $A_1$  that makes it possible to hurt  $A_1$ 's interests by \$5 and that the cost of such a sanction is equivalent to \$5 to either  $A_2$  or  $A_3$ .<sup>2</sup>

These assumptions make it possible to specify for  $A_2$  and  $A_3$  a set of payoffs for the second-order public-good problem posed by the question of how to sanction  $A_1$ . If neither  $A_2$  nor  $A_3$  sanctions, they lose \$1 each (from Table 10.2). If only  $A_2$  sanctions, it costs him \$5 to induce  $A_1$  to contribute (again, the difference between \$8 and \$3). In that case  $A_3$  gains \$4 (the difference between -\$1 and \$3) and ends up with \$3. Similarly, if only  $A_3$  sanctions, it costs him \$5, and  $A_2$  gains \$4. If both sanction, it costs each \$2.50 and each gains \$4, leaving each with a net of \$1.50 after subtracting the cost of the sanction and thus with \$0.50 for the total project.

There are four points of importance to note about this sanctioning problem. First, the structure of payoffs illustrates the fact that the necessity to employ a sanction is costly to the sanctioners. Even though the right to control  $A_1$ 's action is held by  $A_2$  and  $A_3$ , enforcement of that right is costly. The fact that enforcement is possible, at a cost to  $A_2$  and  $A_3$  less than the benefit they can gain from the effects of the sanction in bringing  $A_1$  into line, makes the threat of sanction credible, and the norm viable.

Second, the sanctioning problem involves, as stated earlier, one actor fewer than the original public-good problem. When the original problem involves three actors, the sanctioning problem involves two and can ordinarily be resolved by exchange. With large numbers of actors, of course, the benefit of this size reduction is small.

Third, the sanctioning problem involves a smaller cost to the actors involved than does the original problem. This can be seen in two ways. The difference between the social optimum of a payoff of \$0.50 each and the outcome of -\$1 that results from each taking his individually optimal action is only \$1.50; this difference is \$3 for the original problem. Also, from  $A_3$ 's perspective, if he sanctions, the difference that  $A_2$ 's sanction makes for him is only \$2.50, compared to \$4 in the original problem. Thus the sanctioning problem is less costly than the original problem.

Fourth, if the second-order sanction is a positive one, rewarding the sanctioner, then even though it is less costly than the first-order sanction, it must be

2. That the cost of the sanction to  $A_2$  or  $A_3$  and its benefit to  $A_1$  are equal implies interpersonal comparison of utility, which I have fixed here by fiat by specifying homogeneity among the actors. More generally, however, if  $A_2$  is more powerful than  $A_1$ , he may have available the means to sanction (that is, control over another event), something worth little to him and much to  $A_1$ , making it possible to sanction  $A_1$  with little cost. These differences are discussed in subsequent sections.

provided whenever the right action (sanctioning the initial offender) is taken; a negative sanction must be applied only when the wrong action is taken. If there develops a norm that one must sanction the violator of the initial norm, then the negative second-order sanction for not applying the first-order sanction must be applied only when that sanctioning norm is violated. This cost reduction to norm beneficiaries may give them an interest in establishing a sanctioning norm.

It is now possible to state the second condition for emergence of an effective norm, the condition under which the demand for an effective norm will be satisfied. Stated simply, this condition is that under which the second-order free-rider problem will be overcome by rational holders of a norm. To put it differently, the condition is that under which beneficiaries of a norm, acting rationally, either will be able to share appropriately the costs of sanctioning the target actors or will be able to generate second-order sanctions among the set of beneficiaries that are sufficient to induce effective sanctions of the target actors by one or more of the beneficiaries. This condition depends on the existence of social relationships among the beneficiaries.

### Free Riding and Zeal<sup>3</sup>

The theory developed in this and preceding chapters can be used to solve an empirical puzzle. It is first necessary to restate the free-rider problem: When a number of self-interested persons are interested in the same outcome, which can only be brought about by effort that is more costly than the benefits it would provide to any of them, then, in the absence of explicit organization, there will be a failure to bring about that outcome, even though an appropriate allocation of effort would bring it about at a cost to each which is less than the benefits each would experience.

The puzzle lies in the fact that there are many empirical situations in which just the opposite of free-rider activity seems to occur, even though the circumstances are those in which free riders would be predicted to abound. That is, there is an outcome in which a number of persons are interested, which requires effort whose costs are not fully compensated by the benefits the outcome will bring to any of the persons. Yet in some such situations what is found is the opposite of free riding, that is, an excess of zeal. In the fever of patriotic zeal during wartime, men will volunteer for military service; in the military they will volunteer for front-line duty; and at the front line they will volunteer for dangerous missions. Even among those who are opposed to violence, there are some who will volunteer for front-line duty in providing medical aid to the wounded. In all of these cases the costs that are borne are extreme, including a greatly increased probability of being killed.

Similarly extreme costs are borne by persons in small groups who engage in militant or even terroristic acts on behalf of what they regard as a public good.

3. An earlier version of this section appeared in *Sociological Theory* (Coleman, 1988a).

Examples include the IRA hunger strikers in Northern Ireland, some of whom fasted until death; Mohandas Gandhi and his followers in India, who endured extreme hardship for a cause; the Red Guards in Italy, who engaged in terrorism designed to bring down the system; activists in the PLO in the Middle East; and leaders and activists in the Solidarity movement in Poland. In all these cases a number of persons experienced extreme costs to bring about a result from which they personally could hardly expect to benefit sufficiently to justify those costs.

Another area where free-rider behavior might be expected but zealous activity is often found instead is in team sports. Since the benefits of winning are experienced by all team members, one would expect, by free-rider logic, to find little or no activity by team members. Yet both in practice and in games, team members often work harder than do participants in individual sports (such as track and field events). Even if one accepts the caveat that is often introduced in predictions about free-rider activity—that it does not occur in small groups—this does not explain the higher levels of effort in team sports. It would predict equal levels of effort. What seems instead to occur in team sports is some free-rider activity, that is, some greater amount of loafing than occurs in individual sports, but also zealous activity at a greater level than occurs in individual sports. The overall average level of effort is probably higher in team sports than in individual sports.

### *Rationality of Free Riding and of Zeal*

How can the two phenomena, free riding and zeal, coexist? How can similar situations produce free riding and zeal?

The rationality of free riding is straightforward: If a number of persons' interests are satisfied by the same outcome, and if the benefits that each experiences from his own actions that contribute to the outcome are less than the costs of those actions, he will not contribute if he is rational. If others contribute, he will experience the benefits of the outcome without incurring costs. If others do not contribute, his costs will outweigh his benefits. Yet in much the same situation another rationality leads to zealous activity. If a number of persons' interests are satisfied by the same outcome, then each has an incentive to reward the others for working toward that outcome. Each may in fact find it in his interest to establish a norm toward working for that outcome, with negative sanctions for shirking and positive sanctions for working toward the common goal. If the norm and sanctions do become established, then each person has two sources of satisfaction when he works for the outcome: the objective achievement of his interests through the contribution of his actions toward the outcome, and the rewards provided by the others for helping to achieve that outcome. Thus one's efforts directly help to satisfy one's interests (even if not enough to outweigh the costs of those efforts), and they also bring benefits from others for helping to satisfy *their* interests. The combination of these two benefits can be greater than the costs of the effort one expends.

The rationality of free riding and the rationality of zeal arise under the same structure of interests. This is not the structure of interests that characterizes most situations, where the interests of different persons are complementary and are realized through some kind of social exchange. Nor is it a structure in which interests are opposed, so one person's interests are realized at the expense of another's. Rather, it is a structure of common interests; that is, the interests of all (or at least all in the vicinity) are realized by the same outcome (winning a war or a game, or achieving a political or community goal). It is in these sorts of situations that both free riding and zeal can be found.

How can these two rationalities be made consistent in a way that will allow prediction as to when one or the other will prevail? To answer this requires looking at the similarities and the differences between the rationality of free riding and the rationality of zeal. The rationality of zeal has the same incentive that leads to free riding, but with a second incentive superimposed on the first. The second incentive, however, becomes effective only through an intervening action: encouragement of others, or positive sanctions, which may overcome the deficiency of the first incentive. It is this intervening action that makes the difference between the deficient incentive leading to free riding and the excess incentive leading to zeal. Thus the condition under which free riding occurs and the condition under which zeal is exhibited are delineated by the absence or presence of this intervening activity.

What are the conditions under which the intervening activity is present? When this intervening activity, which I have described as encouragement of others, is examined more closely, it can be seen to be one of a general class of activities that are described as sanctions in enforcement of a norm. (There can be, as I will indicate later, encouragement of others' activity in the absence of a norm, but this can be effective only under special circumstances, which I will specify.) That is, the activity is a certain kind of sanction which *encourages* the action rather than *discouraging* it, and the norm is of a certain kind, one which *prescribes* a certain action rather than proscribing it. But to say that a norm arises under those two conditions is to beg half the question. The existence of externalities is a necessary condition for the existence of an effective norm, but not a sufficient one—if it were, free riding would not exist when actors have common interests.

### *The Closure of Networks and the Emergence of Zeal*

An earlier section showed the importance of social structure in supporting the employment of sanctions. It has also been suggested that it is social structure that can transform free ridership into zeal. How this occurs can be understood by examining differences among social networks, as shown in Figure 11.2. In part a of the figure actors  $A_1$ ,  $A_2$ , and  $A_3$  are not part of the same network. Whatever social relations they have are not with each other. In parts b and c

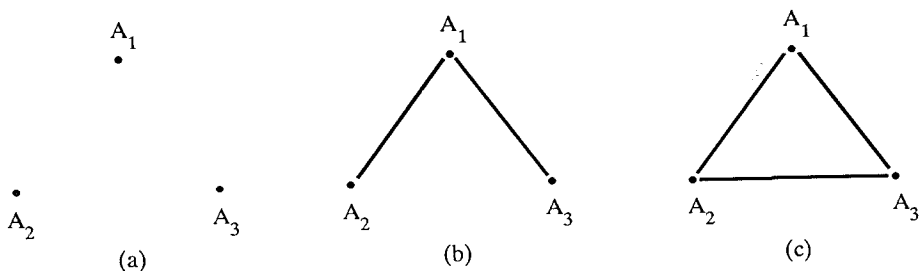


Figure 11.2 Structures of relations among actors that have differing potentials for the growth of zeal.

actor A<sub>1</sub> is connected to actors A<sub>2</sub> and A<sub>3</sub>. In part b actors A<sub>2</sub> and A<sub>3</sub> are not connected, but in part c they are.

Suppose some circumstance arises which fits the incentive structure for both free riding and zeal. That is, there is an activity in which the action of each actor benefits all three, although the benefits to each actor of his own action are insufficient to overcome the costs of that action. The common project introduced in Table 10.2 and discussed in both Chapter 10 and this chapter illustrates such an activity.

In a social structure such as is shown in Figure 11.2(a), each actor has no possibility of influencing the contribution of either of the others. They have no relations, and thus they cannot provide the encouragement or impose the sanctions that will induce the others to contribute. It would be irrational for any of the three to contribute.

In a social structure such as is shown in Figure 11.2(b), matters are more problematic. If A<sub>1</sub> proposes not to contribute, but to free ride on the contributions of A<sub>2</sub> and A<sub>3</sub>, then A<sub>2</sub> would like to induce him to contribute; and in part b, unlike part a, A<sub>2</sub> (and A<sub>3</sub> as well) is in a social-structural position that facilitates this. As shown earlier in connection with Figure 11.1, however, in some circumstances, A<sub>2</sub> cannot do so if he must compensate A<sub>1</sub> in some common medium of exchange in order to encourage him to contribute. There are, of course, public goods for which inducement by A<sub>2</sub> would be sufficient to lead A<sub>1</sub> to contribute while still providing a net benefit to A<sub>2</sub>. (For example, suppose in the three-actor common project A<sub>1</sub>'s contribution of \$9 brought a benefit of \$6 to each. Then A<sub>2</sub> could induce A<sub>1</sub> to contribute by offering him any amount more than \$3 and would benefit if he offered A<sub>1</sub> any amount less than \$6.) It may also be true that A<sub>2</sub> can compensate A<sub>1</sub> with something worth little to him but worth a lot to A<sub>1</sub>. A<sub>2</sub> could offer this to A<sub>1</sub> with a net surplus for both parties. For example, if A<sub>2</sub> is highly respected by A<sub>1</sub>, then A<sub>2</sub>'s gratitude to A<sub>1</sub> for contributing may fulfill this condition.

In the social structure shown in Figure 11.2(c), there is an additional possibility. If a sanctioning of A<sub>1</sub> would cost A<sub>2</sub> (or A<sub>3</sub>) more than he would benefit from it, then with this structure the second-order sanctioning problem could be overcome, as described in the discussion concerning Figure 11.1.

### Is There an Excess of Zeal?

The above discussion simply restates the earlier result that closure of social networks can overcome free-rider activity through the creation of norms and sanctioning systems. The discussion does not indicate, however, how the contributions (of money or effort or time) can be *greater* than would occur if the three actors were each engaged in producing a private good. In the common project the norm brought the contributions back to what they would have been if the return had been purely private goods; that is, if each had received \$12 for his \$9 investment but none from the others' investments. In principle, the existence of a norm with sanctions does what formal organization does in the presence of externalities: It internalizes the externalities.

Social networks, and the norms they facilitate, do more than this, however. Under certain circumstances they generate the excessively zealous activity which indicates not a deficiency of incentives to contribute, but an excess. What leads members of an interconnected group to engage in the opposite of free riding?

Earlier I indicated that if A<sub>2</sub> held something that was worth little to him and much to A<sub>1</sub>, he could induce A<sub>1</sub>'s contribution without loss, even in the social structure of Figure 11.2(b). And in the social structure of Figure 11.2(c), there is the additional possibility of gains from A<sub>3</sub>'s holding something of greater value to A<sub>2</sub> than to himself, which he could use in place of money for the second-order sanctioning of A<sub>2</sub>. The relationships of which social structures are composed contain such possibilities in abundance. An expression of encouragement or gratitude for another's action may cost the actor very little but provide a great reward for the other. The shouts of encouragement to an athlete from his teammates may cost them little but provide him with rewards that lead him to work even harder. Or a girl's smiling at the athlete may cost her little but impel him to new heights of determined effort.

When there are such differentials, the social structure does more than merely internalize externalities. The social system has within it a potential, analogous to the potential in an electrical system. That is, when one actor carries out an action, thus experiencing costs, and others receive the benefits, the return that the actor experiences is not merely those benefits transmitted back to him through the social structure but those benefits amplified by this potential that exists in the structure.<sup>4</sup> Thus when an actor's activity levels off at the point where marginal cost equals marginal return, that point is at a higher level because of the amplified returns he has received, which were in turn produced by the potential that exists in the structure. The potential lies in the difference between the cost to the sanctioner, in each of the relations that transmits rewards back to an actor, and the benefits the sanctioner gains through the increased activity of the actor who is being (positively) sanctioned.

4. This amplification, although it has hardly been studied by social scientists, appears to offer promise for understanding several social-psychological phenomena, such as the intensity of pleasure in new relationships as compared to old ones.

### The Impact of Social Structure

The networks shown in Figure 11.2 exhibit two components of a social structure that are important for the existence of norms which can transform a deficiency of incentive into an excess of incentive. The first of these two components is the existence of social relations between an actor and those for whom he generates externalities. Even if those others are disconnected members of an audience witnessing his actions, their connection to him may mean that they can provide rewards, at a cost to themselves that is below their benefit from his actions, that spur him on to greater efforts. It is for this reason that a performing athlete, musician, or actor may experience far greater motivation than will a book author, who cannot see the reactions of his audience.

The second component of a social structure that is important for the existence of norms is exhibited by the difference between parts b and c of Figure 11.2. This is the closure of the network, the existence of relations between those who experience externalities from another's action. From the example of the common project in conjunction with Figure 11.1, it is clear how these relations can make a difference between a system in which there is a deficient incentive to act and thus a suboptimal level of activity, and a system in which there is an excess incentive to act and thus a manifestation of zeal. More generally, closure of the network gives increased potential for amplifying returns to the actor. Thus a system in which others have connections to the actor may exhibit a strong potential that induces higher levels of activity, but a system that in addition has high closure has an extra potential, due to the benefits that each of those who experience externalities of the actor's action receive from one another. There is an amplification that occurs even before the rewards get back to the actor himself.

### Heroic versus Incremental Sanctioning

Examination of the ways in which norms are characteristically enforced makes it clear that a common mode of sanctioning can be characterized as incremental sanctioning. This is exemplified in unions by "putting in Coventry" scabs and others who seriously violate union norms; that is, other members refuse to speak to the transgressor (Lipset, Trow, and Coleman, 1956). It is also exemplified by the development of a "reputation," which is followed by avoidance or snubbing. In incremental sanctioning the cost incurred by each sanctioner is small, and the effect of each sanctioning is small as well, but the effects are additive, giving a large total effect.

Aesop's fable "The Mice in Council," however, is a reminder that it is not always possible to sanction incrementally. To bell the cat was not an activity that could be engaged in by additive increments. It required what I will call a heroic sanction, that is, a sanction whose total effect occurs through a single actor's action. In the examples of norms and sanctions given earlier in this chapter, the sanctions were imposed by single individuals.

In this section I will use the three-actor common project to examine the different structures of action when sanctions are heroic and when they are incremental, carried out by all the actors in the collectivity other than the one being sanctioned.

Table 11.1 is based on the assumption that a sanction is all or nothing: The sanctioner (assuming only one) must pay the full costs of inducing  $A_1$  to contribute, that is, something worth \$5 to  $A_1$  (which in this homogeneous case is assumed to cost  $A_2$  or  $A_3$  the equivalent of \$5). If the sanctioner is  $A_2$ , he benefits by \$4 from  $A_1$ 's contribution, but his cost of \$5 creates a net loss of \$1. This is a heroic sanction, because  $A_2$  brings about the total effect by his action alone. Unless it is possible to divide the costs with  $A_3$  by both simultaneously sanctioning (see upper-left cell of Table 11.1), he loses \$1 by sanctioning. This \$1 loss may be made up by  $A_3$  as a reward to  $A_2$ , from the \$4  $A_3$  has gained from  $A_2$ 's sanctioning. Thus, when only a heroic sanction is possible, a sequence of two steps is necessary if the heroic sanctioner is to end up with a net benefit from his sanctioning. First  $A_2$  must sanction  $A_1$ , and then  $A_3$  must give a reward to  $A_2$  to compensate for the loss incurred.

If, however, sanctions can be additive in their effects, as empirical evidence suggests they are in many cases,  $A_2$  can bring about a contribution of half of the \$9 from  $A_1$  through a sanctioning action costing  $A_2$  \$2.50 and bringing a benefit of \$2 to  $A_2$  and to  $A_3$ . The net cost to  $A_2$  is only \$0.50. This payoff structure is shown in Table 11.2. There is a prisoner's dilemma structure here, but one with extensive possibilities for mutually beneficial arrangements, due to the disparity between the sanctioner's net loss (only \$0.50) and the other's gain (\$2) from the sanctioning.

This example, however, does not show the differences between heroic and incremental sanctions in a structure of action as clearly as does a case with a

Table 11.2 Payoff matrix for two potential incremental sanctioners in three-person common project.

		$A_3$	
		Sanction	Not sanction
$A_2$	Sanction	.5, .5	-1.5, 1
	Not sanction	1, -1.5	-1, -1

Table 11.3 Net gains for each actor from six-person common project.

Number of contributions	Net gain (in dollars) for—	
	Noncontributors	Contributors
6	—	3
5	10	1
4	8	-1
3	6	-3
2	4	-5
1	2	-7
0	0	—

larger number of actors. Consider the same common project but with six participants, rather than three. Each contributes nothing or \$9, and \$1 is earned for every \$3 contributed; the total product is divided equally among the six. The net gain for each contributor and noncontributor for each configuration of contributions is shown in Table 11.3. The net loss incurred by contributing is no longer \$5, but \$7. (For example, if five are contributing, the noncontributor's gain is \$10. If he contributes, there are six contributing, and he ends up with \$3, making him \$7 worse off.) The net gain experienced by the others from one actor's sanctioning of a noncontributor is no longer \$4, but only \$2. (For example, if the sixth actor does contribute, the others' net gain goes from \$1 to \$3.)

If sanctioning cannot be incremental, the heroic sanctioner in this six-actor project must incur a cost of \$7 to achieve a benefit of only \$2. He has a net loss of \$5, whereas the heroic sanctioner in the three-actor project has a net loss of \$1. Furthermore, this net loss of \$5 cannot be made up by another who benefited without that other's experiencing a net loss, because each actor's benefit from the heroic action is only \$2. Even two others could not provide sufficient rewards to the heroic sanctioner to make his action anything other than foolhardy; if they rewarded him with their gains, he would still have a net loss of \$1. It would take three others, that is, all but one of the four who gained by the heroic sanction, to make up for the heroic sanctioner's net loss.

If the sanctions can be incremental, the degree of exposure of the sanctioner is much less.  $A_2$  as sanctioner, for example, incurs a cost of  $\$7/5$ , or \$1.40, and gains from his sanctioning alone \$0.40 from the incremental contribution made by  $A_1$ , the actor sanctioned (although  $A_2$  gains \$2 altogether if others sanction as well). Thus each actor experiences a net loss of \$1 by sanctioning. It is again possible for this to be made up by a sequence of rewards from others, of \$0.40 each, which would require participation by at least three of the other four. Alternatively, additional incremental sanctions from the others can make up  $A_2$ 's loss, each incremental sanction reducing the loss by \$0.40. If all sanction incrementally,  $A_2$  gains \$0.60. Thus if incremental sanctions are to pay the sanctioner, some prior collective decision that all (or at least many) will sanction is required

(as in the consensus Merry, 1984, p. 279, describes as the second phase of gossip).

For example, suppose all members of a club are expected to clean up after meetings, but one member consistently fails to help. If one person expresses disapproval, this might induce a small effort on the offender's part, but would also have a negative effect on the relationship between these two, an effect that might be more important to the potential sanctioner than the benefit from the offender's efforts. But if all concurred in expressing disapproval, inducing the offender to make his full contribution, the benefits to each would outweigh the costs of each one's worsened relation with the offender.<sup>5</sup>

Returning to the six-actor common project, suppose there is not a binding collective decision, and all but one have sanctioned. Then the sanctions can go one stage deeper. Suppose that  $A_1$  is the noncontributor and  $A_2$  is the nonparticipating in incremental sanctioning. Each of the others has provided an incremental sanction, and  $A_1$  has made four-fifths of his total contribution.  $A_2$ , who is \$1 better off by not sanctioning, can be induced to sanction either by a heroic second-order sanction of \$1, which works out to a net cost of \$0.60 to the second-order sanctioner, or by incremental sanctions of \$0.25, which work out to a net cost of \$0.15 for each of the sanctioners.

The overall difference between the heroic sanction and the incremental sanctions lies in the magnitude of the costs incurred by the sanctioners at every stage. At the first stage in the six-actor project,  $A_2$ , the heroic sanctioner, must incur a net cost of \$5, a loss five times that incurred by each incremental sanctioner. At the second stage, for  $A_3$  to reward the heroic sanctioner alone imposes a net cost of \$3 on  $A_3$ . If sanctioning is incremental, the free-rider problem remains, but at a greatly reduced magnitude. The net cost to each sanctioner is \$1, rather than \$5. If the second-stage sanction (the reward to the incremental sanctioner) is heroic, the second-stage heroic sanctioner incurs a net cost of only \$0.60, rather than \$3. If the second-stage sanction is incremental, the net cost to each of the four sanctioners is only \$0.15.

What this means in practice is that in many circumstances in which heroic sanctions are beyond the resources of any sanctioner, the resources for incremental sanctioning are readily available. These resources may, as indicated earlier in this chapter, be other events which are controlled by each of the potential sanctioners. The values specified above indicate only the costs that sanctioning will impose on the sanctioner, if the cost to the sanctioner is the same as its cost to the person being sanctioned (as in the case of a parent who tells a child that a punishment "hurts me as much as it hurts you"). When the sanctions have the small cost that incremental sanctions come to have in a group of any size, a positive sanction may consist of nothing more than a "credit slip"

5. Empirically the costs might also be reduced, since disapproval from all might lead the offender to accept the collective verdict and not to respond unpleasantly to the members expressing disapproval. In the example of Table 11.3, however, the net gain from incremental sanctioning by all does not depend on such reduced costs.

in the form of gratitude for what the other has done, or a negative sanction may consist of nothing more than a withdrawal of credit in the form of displeasure ("Just wait till you ask me to do something for you!").

Other possibilities exist with incremental sanctioning. If there is some heterogeneity among the potential sanctioners, then, as Chapter 30 shows, the free-rider problem may be overcome at some stage and in any case will constitute less of an obstacle.<sup>6</sup> The complex possibilities that exist can only be alluded to here. In Chapter 30 they are examined in some detail, with the aid of a formal model which facilitates inclusion of the full set of resources available to each actor, not merely those in the common project itself.

I must point out that here the term "heroic" refers to a single sanctioning by one sanctioner sufficient to bring about a noncontributor's contribution. If the set of five contributors (or a large enough subset) can act as a single actor, a single sanction from that set can be sufficient to bring about the contribution and yet bring a net benefit to each. Many communes hold meetings once a week or at some regular interval, at which the whole membership gathers to give self-criticism or hear criticism by others. This phenomenon suggests that, in such settings, this method of sanctioning is easier to organize than either heroic sanctioning or sanctioning by independent increments.

### How Are Sanctions Applied in Society?

The preceding analysis indicated the logical character of the sanctioning problem. The theory construction can be aided by examining some ways that sanctioning is commonly done. This review is not exhaustive, but is only a means of discovering some ways that the second-order public-good problem of sanctioning is solved.

It is useful to begin with some of the examples introduced at the beginning of Chapter 10. In the example involving the dropping of a candy wrapper on a Berlin sidewalk, the sanction was imposed by one woman, without apparent social support. In the example involving the hoarding of tea bags, the sanction was again applied by a single person, another tea drinker in the same organization. Did neither of these sanctioners need any additional support? Was the sanction so near to being costless that it could easily be applied, despite the fact that the direct benefits the sanctioner might experience were uncertain and weak at best, and despite the unpleasantness that might ensue when the disapproval was voiced?

The question cannot of course be answered for these cases in the absence of evidence, but two observations may be made. First, in both cases the sanctioner may paradoxically have depended on some implicit support from the person

6. On the other hand, if the target actor is more powerful than the sanctioner, a sanction that hurts the target actor a little will hurt the sanctioner a lot, thus exacerbating the sanctioning problem.

being sanctioned, that is, the sanctioner may have felt that the person accepted the normative definition of what action is right and recognized that the action carried out was wrong. Second, the sanctioner in either case may have been able to bring up the event in subsequent discussion with others who shared the same opinion or feeling about the event and would provide encouraging comments in support of the disciplining that the sanctioner carried out. If so, this introduces a third actor, comparable to  $A_3$  in Figure 11.2(c), with whom the sanctioner has a relation and in whose approval the sanctioner has some interest. Thus, when it appears that sanctioning is carried out heroically, by a single person without social support, there may in fact be support from other actors on whom the target actor's actions would impose externalities. It is also true that such support for the sanctioning is less costly than the sanctioning itself, and thus has no potential for bringing about unpleasantness, which the original sanctioning might produce for the sanctioner.

This is a general result: Where sanctions are applied in support of a proscriptive norm and are consequently negative sanctions, the second-order public-good problem of providing positive sanctions for the sanctioner is more easily overcome, because positive sanctions incur lower costs than do negative ones.

Another observation concerning these two examples is that whether the sanctioner depended on implicit support from the target actor or on subsequent approval from a third actor, there was an assumption concerning what is right. That is, both mechanisms on which the sanctioner may have depended for support are based on a norm defining what is the right action or (as in these cases) what is the wrong action. The norm, prescribing what is right or proscribing what is wrong, gives a sanctioner some presumption that his action will elicit approval from those who hold the norm. He has a presumptive right to impose the sanction. Thus the existence of a norm provides for a potential sanctioner some expectation of receiving approval from the holders of the norm. However, this expectation is highly contingent on the social relations between the potential sanctioner and other holders of the norm, because establishment of the norm and vesting of the right to sanction can be achieved only by some form of collective decision, implicit or explicit. One difference between the older woman in Berlin and the older woman in New York City may have been that the former spent evenings with others like herself, with whom she could discuss the shortcomings of the younger generation's child-rearing practices and arrive at consensus about what is right, and the latter spent evenings in an apartment alone.

Another example from Chapter 10 involves a different kind of sanctioning than that found in the two examples discussed above. In the case of the poor family among the Sarakatsan nomads in Greece, gossip appears to have played an important part in generating sanctions. Merry (1984), in a review of studies of the role of gossip in social control, suggests that there are three distinct phases.

The first is the circulation of information about an event or action. The second is the formation of some consensus about the moral meaning of that event; how it is to be interpreted, and which rules are to be applied . . . The

third phase is the implementation of the consensus, the transformation of shared opinions into some form of action. This action can range from individual acts of snubbing to collective decisions to expel. (p.279)

Gossip appears to be an important element in the enforcement of norms in many contexts. Why should this be so? It appears that gossip is a means of generating sanctions that could not be applied by individuals in the absence of or before the gossip. If there are three phases of gossip, as Merry suggests, the first two phases appear to be motivated by the potential aid that a consensus provides for the application of a sanction. The consensus either establishes a norm (that is, a definition of what is right or what is wrong, and the assumption of rights to partially control the action) or establishes the application of an existing norm to the action in question.

Each person who has an interest in the maintenance of the norm and the application of sanctions to those who violate it comes thereby to have an interest in the spread of information that can lead to a consensus on legitimate sanctions. This means such a person will be interested in listening to gossip and interested in passing gossip. If the consensus leads to a collective decision to expel the offender, or to cut off communications with him, as occurs in some communes (see Zablocki, 1971), then the second-order public-good problem is overcome. If it does not, the consensus nevertheless provides the basis for support by members of the community of heroic sanctions applied by individuals. This overcomes the second-order problem by a second stage consisting of positive sanctions rewarding the heroic sanctioner. The consensus also increases the effectiveness of incremental sanctions which are sufficiently small that the cost they impose on each sanctioner is minimal—for these sanctions taken together may be very powerful. The combined effect of individual acts of snubbing can mean isolation of the individual, whereas one individual's snub would constitute an ineffective sanction.<sup>7</sup>

Gossip, then, constitutes a device which both aids in establishing a norm and overcomes the second-order public-good problem of sanctioning. It leads to sanctions that may have little cost for the beneficiary of the norm, the one who passes gossip or the one who receives it, and also brings him potential benefits. The benefits lie in the facilitation gossip provides, through the consensus it brings about, for sanctions that might not otherwise be possible. A typical comment that suggests this goes something like this: "Knowing that you feel the same way about what he did gives me the courage to speak to him about it." When such consensus is amplified by many others who feel the same way, the "courage" it provides to each can be great.

But gossip depends on two conditions: First, similar externalities must be experienced by a number of actors, who thereby come to be beneficiaries of

7. Although internal sanctions are not introduced until a later section, it is useful to note here that if the target actor has internalized others' evaluations of him, the very knowledge that gossip is circulating about him can be a strong sanction.

the same norm. If they are all to find benefits in the spread of gossip and the consensus it helps bring about, they must share interests in prescribing or proscribing the action. Second, gossip depends on there being relatively frequent contact among the persons who are affected similarly by an actor's action (and thus are motivated to pass gossip). The benefit that any individual beneficiary can expect to gain from spreading or receiving gossip is relatively small. The costs of doing so must be correspondingly small if persons are to be motivated to spread gossip. Ordinarily these costs are small only if the opportunity to pass gossip is a by-product of relations formed and maintained for other reasons.

This second condition does not involve merely the extent of social relations but the degree to which these relations close in upon themselves as well. Comparison of parts b and c of Figure 11.2 shows that it is the closure of the structure which distinguishes the two. More precisely, what I mean by closure is the frequency of communication between two actors for whom another actor's action has externalities in the same direction.

Empirical studies of gossip confirm the importance of closure. Merry (1984), in a review of work in this area, gives the following generalization:

Gossip flows most readily in highly connected, morally homogeneous social networks, and it is here that its impact is greatest. For gossip to occur, the two participants must know a third party in common. The more mutual friends they have, the more people they can discuss. Every individual is at the center of a network of people he or she knows. The extent to which the members of this network know one another, independent of their relationship to ego, can be described as their "degree of connectedness." Gossip flourishes in close-knit, highly connected social networks but atrophies in loose-knit, unconnected ones.

Moreover, only when the gossipers share moral views is the soil fertile for gossip. The person sharing a juicy tidbit expects the listener to join in condemning that behavior, not to approve it. If only minor differences in norms exist, gossip can forge consensus, but where fundamental ideas of proper behavior differ, gossip will be stunted. (p. 277)

It is evident that both sanctions that have a foundation in gossip and sanctions that do not are more likely to be applied in social structures which exhibit the property of closure than in those which do not. Empirical work shows this, and there are theoretical grounds for such a conclusion. Closure reduces the net costs of applying a sanction, because the consensus that occurs in structures having closure provides the legitimacy (that is, the right) for actors to apply sanctions. This ensures that there will be some compensation (in the form of approval) for the costs incurred in imposing a sanction.

It is clear, however, that gossip itself does not constitute a sanction. Even if gossip analytically has two elements, communication concerning an action and consensus concerning it (the first two phases specified by Merry), neither of these necessarily constitutes a sanction for the target actor. For some persons in



some societies, merely being aware of others talking "behind one's back" constitutes a powerful sanction; gossip can spread, however, without the person who is being gossiped about knowing of it. As indicated earlier, the consensus lowers the costs for any holder of the norm to apply a sanction, but does not ensure that the sanction will be applied.

One social characteristic possessed by a potential target actor is reported by anthropologists and sociologists as reducing the likelihood that sanctions will be imposed: especially high status or power in the social system which contains the norm holders (see, for example, Frankenberg, 1951, p. 156; Bailey, 1971, p. 283). This provides confirmation of the view that the act of sanctioning imposes costs on the sanctioner, since such costs can be expected to be especially high if the target of the sanctioning is someone with whom a continued relation is of special interest to the potential sanctioner. A clear example of this situation is exhibited in the story by Hans Christian Andersen about the emperor who had no clothes. Almost all the inhabitants of the kingdom were so subject to the king's authority that they were afraid to point out the absence of clothes. Only a child was free of such feelings of dependence and thus able to state the truth.

This implies that even a conjoint norm, for which the targets and the holders are the same actors, may be differentially applied because of the varying costs of sanctioning different actors. The consequence is that those actors with greater power in a social system are less constrained by norms than are those with less power. There are, in fact, institutionalized excuses and indulgences available to high-status persons who fail to obey norms. A high-status person may merely be said to be eccentric, whereas the same behavior would bring severe sanctions upon a lower-status person.

At the other extreme there is little cost to applying sanctions to persons whose status is considerably below one's own. Black (1976, p. 57) suggests that legal systems have often meted out harsher punishments to lower-status persons than to higher-status persons for the same violations. Garnsey (1973, pp. 162-164) gives various examples of this in the Roman Empire. If sanctioning of a lower-status person or a member of a small minority can be done without cost, it can be done irresponsibly, as in the case of scapegoating.

#### *A Note on Compliance with Norms*

In general, I will have little to say about compliance with norms, because, in this theory, compliance or noncompliance is merely the result of the application of the principle of maximizing utility under different constraints. There are, however, some empirical results concerning compliance that are closely related to the structural and positional factors discussed above. The points of relevant evidence are these:

1. Powerful persons in a community are not only less likely to be sanctioned, but also less likely to obey the norms than are those of lesser power (Bailey, 1971, p. 20; Starr, 1978, p. 59).

2. Those lowest on the social ladder, although not less likely to be gossiped about or to be negatively sanctioned by others, are less compliant with norms and sanctions than are those above them (Pitt-Rivers, 1971).
3. No matter what degree of closure exists among the holders of a norm, those targets of a norm who have contacts with others outside who are not norm holders, are less likely to be compliant with sanctions (Bott, 1971; Pitt-Rivers, 1971; Merry, 1981).

The first of these empirical results follows from the logic that inhibits those of lesser power in applying sanctions to those of greater power: The latter, recognizing this inhibition, can deviate from norms with less fear of being sanctioned than is true for those below them. Although their position means they have a great deal to lose, it also means they will be less likely to be called to task for deviant actions.

The second empirical result is consistent with the fact that negative sanctions related to social respect can have no effect on those who are at the bottom of the social ladder because they have nothing to lose. This relative disregard of negative sanctions among those without social position should be limited to certain kinds of sanctions, such as disapproval. For other kinds, such as physical punishment, the effectiveness should not be lower among those without social position.

The third empirical result, the imperviousness to norms among those with contacts outside the group of norm holders, is almost transparent. Those with greater mobility can escape sanctioning either physically (as did one young man in Merry's study [1984, p. 292] of an urban neighborhood who simply moved a few blocks away) or psychologically by reducing interest in the community in which sanctions are imposed, increasing them in areas of their lives that lie outside this community.

The third empirical result concerning compliance with norms has implications for the conditions under which norms emerge. Norms emerge as a result of purposive actions on the part of actors who experience externalities from others, potential beneficiaries of the norm. This means that even if there is sufficient closure in the social network of potential beneficiaries, there will be little incentive to bring a norm into being if the potential target actors have sufficient mobility to escape the effect of sanctions. This structural condition exists in some disorganized lower-class urban neighborhoods (as well as among highly mobile high-status persons).

#### *Why Accept a Norm as Legitimate?*

I turn now to a question which introduces more explicitly than before the different interests that arise on the part of different actors when there is some heterogeneity in their control and their interests: Why do persons accept the legitimacy of others' claim to a right to control their action when this acceptance constitutes an immediate disadvantage? It may be that not all such acceptance

can be accounted for by rational choice theory as currently constituted. For example, social psychologists such as Asch (1956) show that a person's very perceptions can be altered by others' reports of different perceptions of the same objects; and Sherif (1936) has shown that judgments may be easily altered by the frame of reference supplied by another. In at least some cases, however, acceptance can be explained by rational choice principles. Although a person may see acceptance of the right of each to partially control the actions of others as being to his *immediate* disadvantage, he may well see it as being to his long-term advantage. If the norm is a conjoint norm, the person may, on some occasions, be in the position of the person expressing this right, that is, in the position of a person affected by another's action of the same type. Acceptance of the legitimacy of others' rights to partially control his action is necessary to establish the norm that gives him a legitimate right to control others' similar actions. Rejection of that legitimacy constitutes a rejection of the norm, an action against the legitimacy of his right on those other occasions. For example, if residents of a dormitory attempt to establish a norm that one cannot use the public telephone for more than 10 minutes if others are waiting, then if one resident of the dormitory rejects the legitimacy of such control, he thereby rejects the norm and cannot claim the right to sanction others when they make long telephone calls. Thus with a conjoint norm a person may rationally accept others' claims to partially control his action, for he stands potentially in the position of sanctioner of others' actions. Acceptance of others' claims is necessary to establish the norm which aids him in controlling their actions which affect him.

It is clear, however, that even for conjoint norms such as that governing telephone calls in a dormitory, there are asymmetries. Some persons, who make many short calls, will often be in the position of sanctioner, and others, who make a few long calls, will often be in the position of target. Depending on a person's perception of how often he will be in these two positions, he may or may not find it to his long-term advantage to accept the norm regarding the length of calls.

Recognition of the fact that persons differ with respect to the relative frequency with which they will be in the position of sanctioner or target can give guidelines for prediction of who will accept a norm as legitimate and who will not. It is less likely to be in the interests of a person who often finds himself the target actor, being constrained by the norm, to accept it as legitimate. The norm is less likely to benefit him in the long run. For example, the claim in a community of a right by the members to constrain—through expressions of disapproval—the bathing suit styles worn at the community swimming pool will be more likely to be regarded as legitimate by the older and less attractive members of the community, who will seldom wear bathing suits that would be challenged, than by the young and beautiful ones, whose swimwear is more likely to be challenged.

It also follows from rationality that those who fail to observe a norm will be less likely to impose the norm on others—for if a person fails to observe the

norm, imposing a sanction on others for not observing it increases the likelihood of being sanctioned oneself. If a girl wears a revealing bathing suit that violates the norm, she is less likely to disapprove of another girl's revealing suit than she would be if she herself wore a conventional suit. If a man is among a group of men, all of whom are wearing ties in a setting where wearing a tie is normatively prescribed, that man's reaction to a newcomer's wearing or not wearing a tie will depend on whether he himself is wearing one.

With a disjoint norm—for which those who are targets of the norm are not beneficiaries of it, such as children whose actions are sanctioned by adults or smokers whose actions are sanctioned by nonsmokers—giving up the right to control one's action must arise from considerations beyond the focal action itself. Giving up that right must be the result of a transaction that can take place because the beneficiaries of the norm (or some subset of them) are powerful; that is, they control some events of interest to the target actors, and they can exchange that control (including exchanges involving threats) to gain rights of control over the focal action.

With a norm that has been internalized, the situation is somewhat similar to that with a conjoint norm. If a person comes to *identify with* a socializing agent, that is, to see his interests as identical to those of the agent, then the claim by that agent of a right to control will be seen as legitimate, because it is a claim deriving from interests the person sees as his own.

### Emergence of Norms about Voting

The act of voting poses a deep and serious problem to students of the rational calculus of behavior. If a voter is viewed as a rational actor who has an interest in the outcome of an election, but for whom the act of voting itself constitutes a small cost of time and effort, then the act of voting does not directly follow, even though his interest in the outcome may be very great. Straightforward considerations will lead the voter to recognize that if there are many others voting, his own vote is very unlikely to affect the outcome. The small cost in time and effort that voting will incur must outweigh the very minor chance that the act of voting will bring gain. As a consequence, a reflective voter must conclude, as he is going to the polling place, that whatever impels him there, it is not the impact of his vote on the outcome.

This sort of analysis has been carried out by many students of the rational calculus of voting. Downs (1957) discusses the problem; Riker and Ordeshook (1973) examine it at length, as do Ferejohn and Fiorina (1974), Margolis (1982), and many others. The empirical fact to be explained is, of course, that although the above considerations seem reasonable, many persons do vote, even in elections involving great numbers of voters. Indeed, even a weaker prediction from rational considerations, to the effect that whatever the level of voting, it should certainly decline as the number of voters increases, does not seem to hold. This problem is so puzzling that it has been described as the paradox of voting (and

has received attention second only to that given Arrow's or Condorcet's paradox). The paradox in this case is not logical, but empirical: Why do so many persons vote when it is clearly irrational to do so?

The problem can be expressed more precisely in simple mathematical form. Suppose a person experiences a certain cost,  $c$ , from the act of voting. And suppose he would experience a benefit,  $b$ , from the election's having the outcome he prefers. Suppose further that he expects that outcome to occur with probability  $p$  if he does not vote and with probability  $p + \Delta p$  if he does vote. Then he can calculate the expected return if he does participate and the expected return if he does not. His expected return if he participates is  $b(p + \Delta p) - c$ . His expected return if he does not participate is  $bp$ . If he is rational, in the usual meaning of the term, he will participate if and only if the first of these two expected returns is greater than the second, that is, if  $b(p + \Delta p) - c$  is greater than  $bp$ . This reduces to  $b\Delta p$  is greater than  $c$ , or  $\Delta p$  is greater than  $c/b$ . That is, he should participate only if the increment in the probability of his desired outcome, due to his vote, is greater than the ratio of the costs of voting to the benefits of the desired outcome.

It is obvious that whenever the number of voters is large,  $\Delta p$  is quite small, so an individual's voting can be explained as rational only if the costs of voting to him ( $c$ ) are nearly zero or the benefits ( $b$ ) are enormous.<sup>8</sup> Few political scientists would suggest that either condition is met for most voters, and thus the puzzle of apparently nonrational voting remains.

Various authors have tried to "rationalize" voting in a variety of ways. One is to assume that the act of voting may not only incur some costs, but also bring some benefits. For example, if voting is highly approved by a person's friends and not voting is disapproved of, the outcome of the election does not have to be of interest to the person in order for him to vote, nor does he have to believe that his vote will affect the outcome. In that case, if the psychic benefits he experiences from approval are  $b^*$  and the costs he experiences from disapproval are  $c^*$ , his expected return from voting is  $b(p + \Delta p) - c + b^*$  and from not voting is  $bp - c^*$ . This changes the inequality that must be satisfied if the person is to vote:  $b\Delta p + b^* + c^*$  must be greater than  $c$ . This inequality may be fulfilled even if he believes his vote will have no effect on the outcome, that is, even if  $\Delta p$  is zero. All that is necessary is that the sum of the psychic benefits from approval for voting and the psychic costs from disapproval for not voting be greater than the direct costs of participation.<sup>9</sup>

8. Another definition of rationality, the minimax regret principle, might also lead a person to vote, as pointed out by Ferejohn and Fiorina (1974). However, this principle can be regarded as rational only in games with a strategic other. The situation under discussion, on the other hand, is a game against nature, a nonstrategic other.

9. The decision to vote can be due to another process discussed in a later section of this chapter and in Chapters 7 and 19, that of coming to identify with a nation or with a political party. This can result in the act of voting itself being rewarding, despite its minimal effect—and even in the absence of approving friends. This analysis of voting behavior would lead to somewhat different predictions than those discussed in this section. For development of a

This explanation of voting has certain virtues. One is that it gives rise to differential predictions about whether persons will vote in different circumstances and can thus be empirically supported or disconfirmed. For example, one prediction would be that persons removed from the company of their friends or persons in the company of those who do not express approval for voting and disapproval for not voting will be much less likely to vote than others.

This explanation has certain unsatisfactory aspects as well. A principal one is that it provides no answer to the question of why others express approval for participation and disapproval for nonparticipation. Those others are presumably subject to the same rational considerations, which should lead them not to vote themselves, as well as not to apply sanctions to others for not voting. Thus the above explanation merely pushes the problem back one step: Why should there be expressions of approval of voting and disapproval of not voting on the part of others?

What is necessary is an explanation of voting that is consistent with the one given above, but takes the further step of explaining why others, despite their similar circumstances, express approval and disapproval. Such an explanation is given in Chapter 30, based on a mathematical model, which provides certain results beyond those possible with a verbal exposition. It is possible, however, to express the general ideas without resort to mathematics.

Assume there is a system of actors, each of whom has an interest (which corresponds to  $b$  in the above discussion) in the outcome of an election and a negative interest in the act of voting itself (which corresponds to  $c$  in the above discussion). The latter may of course be very small relative to the former. Each actor has some small fraction of control over the outcome of the election, through his vote. In such a circumstance each actor's action is of interest to each other actor; that is, the actions have externalities. There will be, in accord with the principle discussed in Chapter 10, a demand for a norm to vote. And, according to the principle presented in this chapter, satisfaction of that demand depends on the existence of social relationships among potential beneficiaries of the norm if the second-order free-rider problem is to be overcome. When those conditions are met, there will be a general transfer of rights of control over the action of voting or not voting, by each to all.

The end result of this transfer of rights of control will be a system in which each actor has given up a large portion of rights of control over his own action (that is, voting or not voting) and has gained in return a small portion of rights of control over the action of each of the other actors. This constitutes, as discussed in an earlier section, the emergence of a conjoint norm. Each actor will exercise the rights gained in the direction of others' voting, but he may exercise the remaining control he holds over his own action against his own voting (because of the cost to him). His exercise of rights of control over others' voting can be enforced through the exhibiting of approval or disapproval. If each actor has given up most of the control over his own action to others, then the potential approval for voting and disapproval for not voting may be sufficient to overcome the costs each will incur by voting.

One might argue that this conjoint norm would be restricted to a subgroup within the system whose members all support the same candidate, since votes for an opposition candidate impose negative externalities. Since persons generally associate with those whose political sentiments are like their own, the application of such a norm (that is, sanctions of approval for voting or disapproval for not voting) will be toward inducing voting among those who will vote in the same direction as oneself. In fact, the observation by students of voting that persons under so-called cross-pressure are less likely to vote may result from differential application of the norm. (See Berelson, Lazarsfeld, and McPhee, 1954, for a discussion of cross-pressures.) The reduced likelihood of voting among persons who are in surroundings different from those to which their background or interests predispose them may result from reduced application of normative sanctions that would lead to voting.

A further implication may be drawn. Since these normative systems are composed of supporters of each candidate, their strength depends on the degree of closure of each. Unless the social networks that link persons together are somewhat distinct, so there is a correlation between the political preferences of friends, these normative systems cannot function. Thus one prediction based on this theory is that the lower that correlation, so that social networks are largely random with respect to political preference, the lower the proportion of the population voting. By a small additional step, a low correlation between social relations and political preference should lower the rate of voting most for those candidates in a minority position in the system under consideration (for example, a city or town).

The central element of this explanation, which was missing in the earlier one, is the giving up of partial rights of control over one's own action and the receiving of partial rights of control over the actions of others, that is, the emergence of a norm. The end result is that control over the voting of each, which was initially held by each alone, becomes widely distributed over the whole set of actors, who exercise that control in the direction of approval for voting and disapproval for not voting—despite the fact that each has some reluctance to vote himself.

### Internalization of Norms

To examine the process by which norms are internalized is to enter waters that are treacherous for a theory grounded in rational choice. Asking the question of how individuals come to have the interests they exhibit is ordinarily not possible in constructing such a theory. Despite the fact that anyone knows, if only through introspection, that interests change, theory based on purposive action must start with purpose, and the theoretical apparatus is applied to realization of that purpose, whatever it may be. A theory based on rational action thus has the same deficiency at the level of the individual (considered as a system) as a theory which begins with societal purposes or social norms has at the level of the social system. This individual-level deficiency is, as I have indicated in Chapter 1, far

less debilitating to social theory than is the deficiency created by a starting with a social purpose or a set of social norms.

It is nevertheless a deficiency, because individual interests do change and individuals do internalize norms. It would be possible to ignore the latter fact and construct a theory that assumed all sanctions were externally imposed. But such a theory would be weaker, because it could not be used to predict the conditions under which and the degree to which norms would be internalized, and less correct, because predictions based on it would fail to take internalization of norms into account. Thus, acknowledging the dangers in going beyond the usual limits of theories based on rational action, I will in this section examine some questions concerning internalization of norms.

I will not use "internalization of a norm" to mean merely accepting a norm as legitimate, accepting the right of others' to partially control one's actions—nor is this the common way in which the term is used. If internalization of a norm were to mean nothing more than that, it would not inhibit the individual's deviant action that is unobserved by others. In this discussion internalization of a norm will mean that an individual comes to have an internal sanctioning system which provides punishment when he carries out an action proscribed by the norm or fails to carry out an action proscribed by the norm.

The question that then arises is that of how the internal sanctioning system comes to be established. This question can be divided into two: First, assuming that internalization of a norm can take place, what are the conditions under which other actors will attempt to bring about an actor's internalization? Second, what are the conditions that will lead an actor to respond to those attempts by internalizing the norm? Only the second of these questions involves going beyond the usual confines of a theory based on rational action, for answering it requires examining why and how individuals change the motivational structure within themselves (motivation is the role that utility plays in a theory of rational action). This second question will be deferred until Chapter 19. The first question, concerning the socializing agents, will be addressed next.

### *Under What Conditions Will Actors Attempt to Bring About Internalization?*

The question to be considered here, then, is why a beneficiary of a norm, or more generally an actor interested in exercising control over another's action, will attempt to establish an internal sanctioning system within a target actor rather than merely using external sanctions as the occasion warrants. The answer is immediately apparent if one takes the perspective of a parent of a small child or that of a police officer coping with crime in a neighborhood. The existence of an internal sanctioning system within the small child or within each person in the neighborhood would make unnecessary the continual external policing of actions. Thus if internalization can be brought about at a sufficiently low cost, it is a more efficient means of social control than is external policing of

actions. The question then becomes, just what are the conditions under which creation of an internal sanctioning system is likely to be more efficient than maintaining external policing of actions?

A first point to recognize is that since norms are devices for controlling actions in the interests of persons other than the actor, internal and external sanctions constitute two forms of policing: internal policing and external policing. The process of creating an internal policing system is part of a broader process which is ordinarily called socialization. It is the installation in the individual of something which may be called a conscience or a superego; I will call it an internal sanctioning system. Persons for whom socialization has been ineffective, in the sense that they have failed to internalize many social norms, are called sociopaths. Those persons and small children are perhaps the most prominent categories of individuals whose internalization of norms is minimal, whose actions are subject to little, if any, internal policing.

Under what conditions is it rational to attempt to bring about an internal sanctioning system? A step toward an answer is to recognize that it is rational to do so when such attempts can be effective at reasonable cost. Thus part of the answer depends on knowledge of the conditions under which attempts at establishing internalization will be effective. Part of the question, however, may be addressed by assuming some receptivity in the individual. In examining this part of the question, I will move back and forth between empirical observation and theoretical argument, using the former to suggest the nature of processes that do not derive directly from a principle of rational action.

First, I must make some general points. Deciding whether internalization of a norm in another actor is rational must involve balancing the cost of bringing about the internalization to a given degree of effectiveness against the discounted future cost of policing to bring about the same degree of compliance, where the degree of compliance is selected by balancing the costs of noncompliance against the costs of sanctioning by the most efficient means (internal or external).

It is also important, of course, to know whether a real distinction exists between different socialization strategies, in that some constitute more nearly external sanctioning and others constitute more nearly internal sanctioning. The literature on socialization practices certainly appears to indicate that the distinction can be made (see, for example, Miller and Swanson, 1958, and Kohn, 1977), and the differences are not esoteric or observable only to initiates. It is relatively easy to distinguish external sanctions from attempts to instill an internal sanctioning system. A parent who slaps a child's hand or withdraws a pleasure to punish an action or who gives the child something, such as candy, to reward an action is employing an external sanction. A parent who shows that an action of a child has hurt or disappointed the parent or who exhibits happiness and expresses love for the child when the child has carried out an action is employing a sanction that both assumes the existence of an internal sanctioning system and attempts to strengthen that system.

Turning back to the question of the conditions under which internalization will be attempted, I begin with an empirical observation. It appears that socializing agents do not simply attempt to inculcate specific norms. A major component of socialization is an attempt to get the individual to *identify* with the socializing agent. This occurs not only in the socialization of children by parents, but in other cases as well. Nation-states use public education as well as various nationalistic events and patriotic propaganda to encourage the individual to identify with the nation, to take its interest as his own. Some business firms (most prominently in Japan but elsewhere as well) attempt to get their employees to identify strongly with the company. Professional schools and graduate schools socialize a candidate into a profession or discipline, leading him to identify himself with the profession or the discipline (to "become a sociologist," for example). Religious orders, the army, and other institutions use various techniques to socialize entering individuals, to give each a new identity. In all these cases it appears that the socialization activities are attempts to create a new self so that the individual's actions will be dictated by the imagined will or purpose of the actor he has identified with: parents, nation-state, company, religious order, profession, or academic discipline. It is then that will which generates the internal sanctions for future actions.

The empirical evidence suggests, then, that a major strategy taken by actors in attempting to internalize norms in another actor is to do so by modifying the self whose interests the actor will attempt to maximize by his actions. This is an indirect strategy, for it does not attempt to inculcate directly the belief that certain actions are right and others wrong. The strategy is to change the self and let the new self decide what is right and what is wrong (for example, by imagining what one's mother would say about a particular action).

Another perspective on this strategy can be gained by examining the problem that exists for corporate actors in motivating agents to act in their interests. This was described in Chapter 7 as the problem a principal has in getting an agent to act in the principal's interest. A major question the principal must ask is whether to engage in external policing, that is, supervision, or to attempt to bring about internal policing by the agent himself. The means by which the latter is attempted are numerous. Those which use directly economic incentives include stock options, stock ownership, piecework, performance bonuses, commissions, and other means. The first two of these are designed to make the agent's interest largely coincide with that of the principal (in that the agent's interest is satisfied through the outcome or product that also satisfies the principal's interest). The other three tie the agent's interest to what he does for the principal. In addition to economic incentives firms use other means which, like stock ownership, lead agents to identify directly with the firm. Collective activities such as sports clubs, company outings, and various company-sponsored entertainments exemplify these. Creating an expectation of long-term employment with the company is another means.

These actions on the part of corporate actors to induce identification in their

agents, or to induce their agents to act in a way that will be best for the corporate actor, appear to apply a strategy similar to that used by actors engaged in socialization of individuals. The strategy is to attempt to align the agent's interests so fully with those of the principal that the agent's self-interest comes to coincide with the principal's interest. Some of the means used appear to carry out this alignment in a more fundamental way than do others. The strategy of the socializing agent and the strategy of the principal appear to be similar in that neither is an attempt to create internal sanctions for a particular action, that is, to inculcate a norm with respect to one action, a norm with respect to a second action, and so on, on a case-by-case basis. Each strategy goes a level deeper, by modifying the interests of the individual being socialized or of the agent of the principal. The socialization strategy, as well as certain of the economic and noneconomic means used by corporate actors, appears to go yet another level deeper, by creating a new self which takes another's imagined will as the basis for action. The similarity of these two strategies, one bringing about changes within an individual actor and the other bringing about changes within a corporate actor (also in part by bringing about changes within individual agents), suggests that a global strategy may in some circumstances be more efficient than a case-by-case strategy. Assuming that this is so, it is possible to ask about the conditions that would affect the efficiency of internalization.

First, internalization becomes increasingly efficient the greater the number of different types of actions that the socializing actor, such as a parent, wants to control using norms. The basic process of creating identification with the socializing agent constitutes in effect a capital cost, and there is an additional marginal cost for each different action that is to be subject to a normative constraint. The capital cost seems to be the larger by far of these two types of costs, so the total cost to the sanctioning agent of subjecting a large number of actions to normative constraints is not much greater than the total cost of doing so for a small number of actions. Thus, for example, once a mother has brought about a condition in which her daughter has internalized the mother's wishes, the mother can extend the number of prescriptions and proscriptions contained in her wishes without incurring great cost.

This implies that authority systems that, like religious orders, aim to penetrate all aspects of the member's or subordinate's life will be more likely to attempt to bring about identification through creation of a new self than will authority systems that attempt to exercise less broad social control. A similar but less obvious prediction is that parents whose desired scope of authority over their children is broad will make more use of internalization than will parents whose desired scope of authority is narrower. More particularly, parents who have an egalitarian ideology about raising children, leading to a laissez-faire parental authority system (few prescriptions and proscriptions), will find it less efficient to bring about identification and instill internalization of norms in a child than will parents whose desired scope of authority is broader and more rigid. Thus children of egalitarian parents should show less internalization of norms, or be

more nearly sociopathic, than children of authoritarian parents. This prediction counters the commonsense assumption that "enlightened" parents are both more egalitarian and more likely to use internal sanctions than external ones; that assumption leads to the prediction that the correlation between parental egalitarianism and use of internal sanctions is positive.

A second point is that parents (or other actors) who are in a position to establish an internal sanctioning system do not reap all the benefits from it. Parents must pay the costs of internalization, but others will experience some of the future benefits. It is true that parents experience some benefits during the period the child is at home. Since these are only a fraction of the benefits, however, there is an expected underinvestment in internalization from the perspective of the total set of benefits to others that internalization will bring about (reduced policing and fewer negative externalities). This underinvestment should be especially great for internalization of norms which have least to do with a child's actions in the home and are primarily concerned with actions toward others later in life.

This underinvestment in internalization is comparable to and derives largely from the same interests as business firms' underinvestment in human capital (see Becker, 1976, for a discussion of underinvestment in human capital). A firm recovers only a portion of the investment in human capital, depending on the length of tenure of the employee.<sup>10</sup> In the literature on human capital, a distinction is made between specific and general human capital. A firm is able to capture all the benefits of investments in specific human capital (firm-specific knowledge or skills) but not able to capture the benefits of general human capital, which can be used in other firms to which an employee might move. This distinction is similar to the distinction between norms covering a child's actions in the home and norms covering actions outside the home or later in life. The implication of this distinction is that there will be less underinvestment in internalization of a norm prescribing honesty, which is part of the child's actions in the home, than in internalization of a norm prescribing fairness to peers, a trait manifested largely outside the home, on the playground, and in school. There should be even greater underinvestment in internalization of norms proscribing sharp practices, which are exhibited primarily in business later in life.

A further prediction is that underinvestment in internalization of norms should be greater in cultures or settings in which children leave home at a younger age. In such settings a parent, engaging in the socialization of children, will experience a smaller fraction of its total benefits and thus will find the investment in internalization to have a lower payoff. One specific prediction is that in modern societies, where the typical household is two-generational, there should be greater underinvestment than in traditional societies, where there tend to be

10. There is at least this difference: The investment in human capital makes the employee more valuable in the labor market, increasing the chance that he will leave the firm and further decreasing the firm's return on its investment. There is no analogous situation for the socialization of children.

three-generational households or extended families. A second specific prediction is that an increase in divorce rates should decrease the investment in internalization, and children of divorced parents should show a lower degree of internalization. Either divorced parent expects to spend less time with a child and will thus find it less costly to use external sanctions in cases where, if there were a more extended payoff period, the creation of internal sanctions would be more efficient.

A third point is that parents can increase the return on their investment in internalization of norms by identifying with the child and continuing to inform themselves about the child's actions later in life. Such identification, together with information about the child's actions that accord with a parent's wishes and gain approval from others, can bring satisfaction to the parent and thus make an investment in internalization profitable. This would lead to the prediction that the use of internal rather than external sanctions by a parent will be greater if the parent expects that the future associates of the child will hold the same values as the parent does (for instance, as occurs in a stable society).

A specific prediction that follows from this general point is that persons in America and Europe who grew up in the 1960s and experienced a great gap in values between themselves and their parents can be expected to find identification with their own children as they grow up a less profitable investment, and thus a weaker spur toward increasing the investment in creating internal sanctions in their children.

A fourth point concerns different families in the same society. Some families have a strong interest in their status in the community and see family members' actions, throughout life, as affecting that status. Other families have little status in the community, and thus little to lose by the deviant behavior of family members throughout life. Parents in the latter type of family can be expected to seriously underinvest in creating internal sanctions, and parents in the former type can be expected to invest much more heavily in creating such sanctions. Studies of socialization practices of different social groups are quite consistent with this prediction, showing that the lower the social status, the less internal sanctions are used (see Kohn, 1977). Further predictions could be tested as well. Any aspect of social structure which reduces the degree to which the child's later actions will benefit or harm the family's interests (such as residence in a more anonymous urban setting as compared to a small-town setting, or geographic mobility and discontinuity in family life) should weaken the relation between the family's interest in its status and the degree to which socialization practices incorporate internal sanctions. Thus, as these conditions proliferate, families of the same social status will use internal sanctions less often and external sanctions more often; persons in future generations will be decreasingly socialized.

A fifth point makes use of a result from the literature on the economics of agency. The efficiency of supervision (that is, external policing), relative to some incentive system that provides internal policing, is reduced as the actions subject

to observation become more costly to observe. Thus a worker in a cottage industry is more likely to be paid by the piece and less likely to have supervision than is a factory worker making the same product, and an outside sales representative is likely to receive a higher fraction of compensation as commissions and have less supervision than is an inside salesperson or sales clerk. This principle leads to the prediction that the trait of honesty, which is often difficult to observe, will be more likely to be internalized than will the trait of cleanliness or orderliness, both of which are more easily observed and thus more readily subject to external sanctions.

These five points concerning conditions which will lead parents to instill internal sanctions rather than using external policing lead to predictions that are not at all trivial. Certain of the predictions, if they are borne out by research, have strong implications for social control in the future, for they point to decreasing levels of internalization of norms among future generations, assuming that the family continues as the principal agency of socialization. This implies that either use of external policing systems will increase or there will be lower levels of social control.<sup>11</sup>

11. These predictions, and the points which generate them, make evident the naiveté of certain assumptions, such as "Increasing levels of education will increase the viability of democracy" or "Increasing enlightenment through education will result in parents raising better-socialized children."