

SOCIAL THEORY

ROOTS AND BRANCHES

Readings

Second Edition

Peter Kivisto
Augustana College

2003



Roxbury Publishing Company
Los Angeles, California

X. Phenomenology and Ethnomethodology

42. Indirect Social Relationships 262
Alfred Schutz
In this essay Schutz describes the characteristic features of social relationships other than those that are direct or face-to-face.
43. Rules of Conversational Sequence 268
Harvey Sacks
In an analysis of telephone conversations, Sacks offers an ethnomethodological account of the rules or "ethnomethods" people employ in the process of achieving orderly and stable interactional exchanges.
44. The Rational Properties of Scientific and Common Sense Activities. 275
Harold Garfinkel
Garfinkel, who coined the term "ethnomethodology," distinguishes the alternative interpretive frames used in the construction of "common sense" versus scientific rationalities.
45. Partnership 284
Aron Gurwitsch
Gurwitsch offers a phenomenological analysis of the situational bonds that emerge when two people engage in situated and coordinated actions that come to constitute a partnership.

XI. Exchange Theory and Rational Choice Theory

46. Social Behavior as Exchange 295
George C. Homans
One of the key figures associated with the development of exchange theory, Homans sketches an outline of an exchange paradigm that in its most elementary form seeks to explain social behavior in terms of costs and rewards.
47. Power-Dependence Relations* 305
Richard M. Emerson
In this classic essay in the exchange theory tradition, Emerson seeks to provide a link among the concepts of "power," "authority," "legitimacy," and "structure" by articulating a view of power that emphasizes its relational character. Power is thus conceived in terms of "ties of mutual dependence."

*New in the Second Edition.

48. Human Capital and Social Capital 317
James S. Coleman
 Coleman was a major spokesperson for rational choice theory, an approach wherein actors and resources constitute two central elementary concepts. Here he discusses two important types of resources: human capital and social capital.
49. The Emergence of Cooperative Social Institutions 325
Michael Hechter
 From a rational choice perspective, Hechter develops a "solidaristic" explanation for how social institutions are created; key to his argument is an attempt to wrestle with the "free-rider problem."
50. Formulation of Exchange Theory. 334
Peter Blau
 Blau presents an exchange theoretical account that, in contrast to the earlier version of Homans, attempts to focus on social structure and appreciates the differences between the micro-level and the macro-level.

XII. Feminist Theory

51. Subversive Bodily Acts. 342
Judith Butler
 Butler's anti-essentialist feminist theory seeks to question the generally taken-for-granted nature of gender categories and calls for a view that appreciates their inherent fluidity and variability.
52. Toward an Afrocentric Feminist Epistemology 350
Patricia Hill Collins
 Collins endeavors to link an Afrocentric perspective to feminist theory, contending that a black feminist epistemology begins with an appreciation of the concrete experiences of daily life.
53. Sociology From Women's Experience: A Reaffirmation. . . . 360
Dorothy E. Smith
 Smith makes a case for a theoretical perspective that begins with women's concrete experiences; here she discusses her understanding of the notion of "standpoint."
54. Rethinking Freud on Women 371
Nancy Chodorow
 Unlike many feminists who are highly critical of Freud, Chodorow finds value in Freud's work and calls for a reconsideration of psychoanalytic theory.

XI. Exchange Theory and Rational Choice Theory

46

Social Behavior as Exchange

George C. Homans

George C. Homans (1910–1989), a Boston Brahmin, was one of the key figures associated with the development of modern exchange theory, which he intended as an alternative to the grand sociological theorizing of his Harvard colleague, Talcott Parsons. Homans argues that sociological theory ought to be grounded in neoclassical economic theory and in behaviorist psychology, associated with figures such as B. F. Skinner. As such he advocates a form of psychological reductionism. In this essay, published in 1958, Homans sketches an outline of an exchange paradigm, which in its most elementary form seeks to explain social behavior in terms of costs and rewards. He sees social exchange as offering sociology a set of general propositions that, in explaining human behavior, constitute an essential starting point for examining issues related to social structure.

The Problems of Small-Group Research

This essay will hope to honor the memory of George Simmel in two different ways. So far as it pretends to be suggestive rather than

conclusive, its tone will be Simmel's; and its subject, too, will be one of his. Because Simmel, in essays such as those on sociability, games, coquetry, and conversation, was an analyst of elementary social behavior, we call him an ancestor of what is known today as small-group research. For what we are really studying in small groups is elementary social behavior: what happens when two or three persons are in a position to influence one another, the sort of thing of which those massive structures called "classes," "firms," "communities," and "societies" must ultimately be composed.

As I survey small-group research today, I feel that, apart from just keeping on with it, three sorts of things need to be done. The first is to show the relation between the results of experimental work done under laboratory conditions and the results of *quasi*-anthropological field research on what those of us who do it are pleased to call "real-life" groups in industry and elsewhere. If the experimental work has anything to do with real life—and I am persuaded that it has everything to do—its propositions cannot be inconsistent with those discovered through the field work. But the consistency has not yet been demonstrated in any systematic way.

The second job is to pull together in some set of general propositions the actual results from the laboratory and from the field, of work on small groups—propositions that at least sum up, to an approximation, what happens in elementary social behavior, even though we may not be able to explain why the propositions should take the form they do. A great amount of work has been done, and more appears every day, but what it all amounts to in the shape of a set of propositions from which, under specified condi-

tions, many of the observational results might be derived, is not at all clear—and yet to state such a set is the first aim of science.

The third job is to begin to show how the propositions that empirically hold good in small groups may be derived from some set of still more general propositions. “Still more general” means only that empirical propositions other than ours may also be derived from the set. This derivation would constitute the explanatory stage in the science of elementary social behavior, for explanation is derivation.¹ (I myself suspect that the more general set will turn out to contain the propositions of behavioral psychology. I hold myself to be an “ultimate psychological reductionist,” but I cannot know that I am right so long as the reduction has not been carried out.)

I have come to think that all three of these jobs would be furthered by our adopting the view that interaction between persons is an exchange of goods, material and non-material. This is one of the oldest theories of social behavior, and one that we still use every day to interpret our own behavior, as when we say, “I found so-and-so rewarding”; or “I got a great deal out of him”; or, even, “Talking with him took a great deal out of me.” But, perhaps just because it is so obvious, this view has been much neglected by social scientists. So far as I know, the only theoretical work that makes explicit use of it is Marcel Mauss’s *Essai sur le don*, published in 1925, which is ancient as social science goes.² It may be that the tradition of neglect is now changing and that, for instance, the psychologists who interpret behavior in terms of transactions may be coming back to something of the sort I have in mind.³

An incidental advantage of an exchange theory is that it might bring sociology closer to economics—that science of man most advanced, most capable of application, and, intellectually, most isolated. Economics studies exchange carried out under special circumstances and with a most useful built-in numerical measure of value. What are the laws of the general phenomenon of which economic behavior is one class?

In what follows I shall suggest some reasons for the usefulness of a theory of social

behavior as exchange and suggest the nature of the propositions such a theory might contain.

An Exchange Paradigm

I start with the link to behavioral psychology and the kind of statement it makes about the behavior of an experimental animal such as the pigeon.⁴ As a pigeon explores its cage in the laboratory, it happens to peck a target, whereupon the psychologist feeds it corn. The evidence is that it will peck the target again; it has learned the behavior, or, as my friend Skinner says, the behavior has been reinforced, and the pigeon has undergone *operant conditioning*. This kind of psychologist is not interested in how the behavior was learned: “learning theory” is a poor name for his field. Instead, he is interested in what determines changes in the rate of emission of learned behavior, whether pecks at a target or something else.

The more hungry the pigeon, the less corn or other food it has gotten in the recent past, the more often it will peck. By the same token, if the behavior is often reinforced, if the pigeon is given much corn every time it pecks, the rate of emission will fall off as the pigeon gets *satiated*. If, on the other hand, the behavior is not reinforced at all, then, too, its rate of emission will tend to fall off, though a long time may pass before it stops altogether, before it is *extinguished*. In the emission of many kinds of behavior the pigeon incurs *aversive stimulation*, or what I shall call “cost” for short, and this, too, will lead in time to a decrease in the emission rate. Fatigue is an example of a “cost.” Extinction, satiation, and cost, by decreasing the rate of emission of a particular kind of behavior, render more probable the emission of some other kind of behavior, including doing nothing. I shall only add that even a hard-boiled psychologist puts “emotional” behavior, as well as such things as pecking, among the unconditioned responses that may be reinforced in operant conditioning. As a statement of the propositions of behavioral psychology, the foregoing is, of course, inadequate for any purpose except my present one.

We may look on the pigeon as engaged in an exchange—pecks for corn—with the psychologist, but let us not dwell upon that, for the behavior of the pigeon hardly determines the behavior of the psychologist at all. Let us turn to a situation where the exchange is real, that is, where the determination is mutual. Suppose we are dealing with two men. Each is emitting behavior reinforced to some degree by the behavior of the other. How it was in the past that each learned the behavior he emits and how he learned to find the other's behavior reinforcing we are not concerned with. It is enough that each does find the other's behavior reinforcing, and I shall call the reinforcers—the equivalent of the pigeon's corn—*values*, for this, I think, is what we mean by this term. As he emits behavior, each man may incur costs, and each man has more than one course of behavior open to him.

This seems to me the paradigm of elementary social behavior, and the problem of the elementary sociologist is to state propositions relating the variations in the values and costs of each man to his frequency distribution of behavior among alternatives, where the values (in the mathematical sense) taken by these variables for one man determine in part their values for the other.⁵

I see no reason to believe that the propositions of behavioral psychology do not apply to this situation, though the complexity of their implications in the concrete case may be great indeed. In particular, we must suppose that, with men as with pigeons, an increase in extinction, satiation, or aversive stimulation of any one kind of behavior will increase the probability of emission of some other kind. The problem is not, as it is often stated, merely, what a man's values are, what he has learned in the past to find reinforcement but how much of any one value his behavior is getting him now. The more he gets, the less valuable any further unit of that value is to him, and the less often he will emit behavior reinforced by it.

The Influence Process

We do not, I think, possess the kind of studies of two-person interaction that would

either bear out these propositions or fail to do so. But we do have studies of larger numbers of persons that suggest that they may apply, notably the studies by Festinger, Schachter, Back, and their associates on the dynamics of influence. One of the variables they work with they call *cohesiveness*, defined as anything that attracts people to take part in a group. Cohesiveness is a value variable; it refers to the degree of reinforcement people find in the activities of the group. Festinger and his colleagues consider two kinds of reinforcing activity: the symbolic behavior we call "social approval" (sentiment) and activity valuable in other ways, such as doing something interesting.

The other variable they work with they call *communication* and others call *interaction*. This is a frequency variable: it is a measure of the frequency of emission of valuable and costly verbal behavior. We must bear in mind that, in general, the one kind of variable is a function of the other.

Festinger and his co-workers show that the more cohesive a group is, that is, the more valuable the sentiment or activity the members exchange with one another, the greater the average frequency of interaction of the members.⁶ With men, as with pigeons, the greater the reinforcement, the more often is the reinforced behavior emitted. The more cohesive a group, too, the greater the change that members can produce in the behavior of other members in the direction of rendering these activities more valuable.⁷ That is, the more valuable the activities that members get, the more valuable those that they must give. For if a person is emitting behavior of a certain kind, and other people do not find it particularly rewarding, these others will suffer their own production of sentiment and activity, in time, to fall off. But perhaps the first person has found their sentiment and activity rewarding, and, if he is to keep on getting them, he must make his own behavior more valuable to the others. In short, the propositions of behavioral psychology imply a tendency toward a certain proportionality between the value to others of the behavior a man gives them and the value to him of the behavior they give him.⁸

Schachter also studied the behavior of members of a group toward two kinds of other members, “conformers” and “deviates.”⁹ I assume that conformers are people whose activity the other members find valuable. For conformity is behavior that coincides to a degree with some group standard or norm, and the only meaning I can assign to *norm* is “a verbal description of behavior that many members find it valuable for the actual behavior of themselves and others to conform to.” By the same token, a deviate is a member whose behavior is not particularly valuable. Now Schachter shows that, as the members of a group come to see another member as a deviate, their interaction with him—communication addressed to getting him to change his behavior—goes up, the faster the more cohesive the group. The members need not talk to the other conformers so much; they are relatively satiated by the conformers’ behavior: they have gotten what they want out of them. But if the deviate, by failing to change his behavior, fails to reinforce the members, they start to withhold social approval from him: the deviate gets low sociometric choice at the end of the experiment. And in the most cohesive groups—those Schachter calls “high cohesive-relevant”—interaction with the deviate also falls off in the end and is lowest among those members that rejected him most strongly, as if they had given him up as a bad job. But how plonking can we get? These findings are utterly in line with everyday experience.

Practical Equilibrium

At the beginning of this paper I suggested that one of the tasks of small-group research was to show the relation between the results of experimental work done under laboratory conditions and the results of field research on real-life small groups. Now the latter often appear to be in practical equilibrium, and by this I mean nothing fancy. I do not mean that all real-life groups are in equilibrium. I certainly do not mean that all groups must tend to equilibrium. I do not mean that groups have built-in antidotes to change: there is no homeostasis here. I do not mean

that we assume equilibrium. I mean only that we sometimes *observe* it, that for the time we are with a group—and it is often short—there is no great change in the values of the variables we choose to measure. If, for instance, person A is interacting with B more than with C both at the beginning and at the end of the study, then at least by this crude measure the group is in equilibrium.

Many of the Festinger-Schachter studies are experimental, and their propositions about the process of influence seem to me to imply the kind of proposition that empirically holds good of real-life groups in practical equilibrium. For instance, Festinger *et al.* find that, the more cohesive a group is, the greater the change that members can produce in the behavior of other members. If the influence is exerted in the direction of conformity to group norms, then, when the process of influence has accomplished all the change of which it is capable, the proposition should hold good that, the more cohesive a group is, the larger the number of members that conform to its norms. And it does hold good.¹⁰

Again, Schachter found, in the experiment I summarized above, that in the most cohesive groups and at the end, when the effort to influence the deviate had failed, members interacted little with the deviate and gave him little in the way of sociometric choice. Now two of the propositions that hold good most often of real-life groups in practical equilibrium are precisely that the more closely a member’s activity conforms to the norms the more interaction he receives from other members and the more liking choices he gets from them too. From these main propositions a number of others may be derived that also hold good.¹¹

Yet we must ever remember that the truth of the proposition linking conformity to liking may on occasion be masked by the truth of other propositions. If, for instance, the man that conforms to the norms most closely also exerts some authority over the group, this may render liking for him somewhat less than it might otherwise have been.¹²

Be that as it may, I suggest that the laboratory experiments on influence imply propo-

sitions about the behavior of members of small groups, when the process of influence has worked itself out, that are identical with propositions that hold good of real-life groups in equilibrium. This is hardly surprising if all we mean by equilibrium is that all the change of which the system is, under present conditions, capable has been effected, so that no further change occurs. Nor would this be the first time that statics has turned out to be a special case of dynamics.

Profit and Social Control

Though I have treated equilibrium as an observed fact, it is a fact that cries for explanation. I shall not, as structural-functional sociologists do, use an assumed equilibrium as a means of explaining or trying to explain, why the other features of a social system should be what they are. Rather, I shall take practical equilibrium as something that is itself to be explained by the other features of the system.

If every member of a group emits at the end of, and during, a period of time much the same kinds of behavior and in much the same frequencies as he did at the beginning, the group is for that period in equilibrium. Let us then ask why any one member's behavior should persist. Suppose he is emitting behavior of value A_1 . Why does he not let his behavior get worse (less valuable or reinforcing to the others) until it stands at $A_1 - \Delta A$? True, the sentiments expressed by others toward him are apt to decline in value (become less reinforcing to him), so that what he gets from them may be $S_1 - \Delta S$. But it is conceivable that, since most activity carries cost, a decline in the value of what he emits will mean a reduction in cost to him that more than offsets his losses in sentiment. Where, then, does he stabilize his behavior? This is the problem of social control.¹³

Mankind has always assumed that a person stabilizes his behavior, at least in the short run, at the point where he is doing the best he can for himself under the circumstances, though his best may not be a "rational" best, and what he can do may not be at all easy to specify, except that he is not apt to think like one of the theoretical antagonists

in the *Theory of Games*. Before a sociologist rejects this answer out of hand for its horrid profit-seeking implications, he will do well to ask himself if he can offer any other answer to the question posed. I think he will find that he cannot. Yet experiments designed to test the truth of the answer are extraordinarily rare.

I shall review one that seems to me to provide a little support for the theory, though it was not meant to do so. The experiment is reported by H. B. Gerard, a member of the Festinger-Schachter team, under the title "The Anchorage of Opinions in Face-to-Face Groups."¹⁴ The experimenter formed artificial groups whose members met to discuss a case in industrial relations and to express their opinions about its probable outcome. The groups were of two kinds: high-attraction groups, whose members were told that they would like one another very much, and low-attraction groups, whose members were told that they would not find one another particularly likable.

At a later time the experimenter called the members in separately, asked them again to express their opinions on the outcome of the case, and counted the number that had changed their opinions to bring them into accord with those of other members of their groups. At the same time, a paid participant entered into a further discussion of the case with each member, always taking, on the probable outcome of the case, a position opposed to that taken by the bulk of the other members of the group to which the person belonged. The experimenter counted the number of persons shifting toward the opinion of the paid participant.

The experiment had many interesting results, from which I choose only those summed up in Tables 46-1 and 46-2. The three different agreement classes are made up of people who, at the original sessions, expressed different degrees of agreement with the opinions of other members of their groups. And the figure 44, for instance, means that, of all members of high-attraction groups whose initial opinions were strongly in disagreement with those of other members, 44 per cent shifted their opinion later toward that of others.

Table 46-1

<i>Percentage of Subjects Changing Toward Someone in the Group</i>			
	Agree- ment	Mild Disagree- ment	Strong Disagree- ment
High Attraction. . . .	0	12	44
Low Attraction. . . .	0	15	9

Table 46-2

<i>Percentage of Subjects Changing Toward the Paid Participant</i>			
	Agree- ment	Mild Disagree- ment	Strong Disagree- ment
High Attraction. . . .	7	13	25
Low Attraction. . . .	20	38	8

In these results the experimenter seems to have been interested only in the differences in the sums of the rows, which show that there is more shifting toward the group, and less shifting toward the paid participant, in the high-attraction than in the low-attraction condition. This is in line with a proposition suggested earlier. If you think that the members of a group can give you much—in this case, liking—you are apt to give them much—in this case, a change to an opinion in accordance with their views—or you will not get the liking. And, by the same token, if the group can give you little of value, you will not be ready to give it much of value. Indeed, you may change your opinion so as to depart from agreement even further, to move, that is, toward the view held by the paid participant.

So far so good, but, when I first scanned these tables, I was less struck by the difference between them than by their similarity. The same classes of people in both tables showed much the same relative propensities to change their opinions, no matter whether the change was toward the group or toward the paid participant. We see, for instance, that those who change least are the high-attraction, agreement people and the low-attraction, strong-disagreement ones. And those who change most are the high-attrac-

tion, strong-disagreement people and the low-attraction, mild-disagreement ones.

How am I to interpret these particular results? Since the experimenter did not discuss them, I am free to offer my own explanation. The behavior emitted by the subjects is opinion and changes in opinion. For this behavior they have learned to expect two possible kinds of reinforcement. Agreement with the group gets the subject favorable sentiment (acceptance) from it, and the experiment was designed to give this reinforcement a higher value in the high-attraction condition than in the low-attraction one. The second kind of possible reinforcement is what I shall call the “maintenance of one’s personal integrity,” which a subject gets by sticking to his own opinion in the face of disagreement with the group. The experimenter does not mention this reward, but I cannot make sense of the results without something much like it. In different degrees for different subjects, depending on their initial positions, these rewards are in competition with one another: they are alternatives. They are not absolutely scarce goods, but some persons cannot get both at once.

Since the rewards are alternatives, let me introduce a familiar assumption from economics—that the cost of a particular course of action is the equivalent of the foregone value of an alternative¹⁵—and then add the definition: Profit = Reward – Cost.

Now consider the persons in the corresponding cells of the two tables. The behavior of the high-attraction, agreement people gets them much in the way of acceptance by the group, and for it they must give up little in the way of personal integrity, for their views are from the start in accord with those of the group. Their profit is high, and they are not prone to change their behavior. The low-attraction, strong-disagreement people are getting much in integrity and they are not giving up for it much in valuable acceptance, for they are members of low-attraction groups. Reward less cost is high for them, too, and they change little. The high-attraction, strong-disagreement people are getting much in the way of integrity, but their costs in doing so are high, too, for they are in high-attraction groups and thus foregoing

much valuable acceptance by the group. Their profit is low, and they are very apt to change, either toward the group or toward the paid participant, from whom they think, perhaps, they will get some acceptance while maintaining some integrity. The low-attraction, mild-disagreement people do not get much in the way of integrity, for they are only in mild disagreement with the group, but neither are they giving up much in acceptance, for they are members of low-attraction groups. Their rewards are low; their costs are low too, and their profit—the difference between the two—is also low. In their low profit they resemble the high-attraction, strong-disagreement people, and, like them, they are prone to change their opinions, in this case, more toward the paid participant. The subjects in the other two cells, who have medium profits, display medium propensities to change.

If we define profit as reward less cost, and if cost is value foregone, I suggest that we have here some evidence for the proposition that change in behavior is greatest when perceived profit is least. This constitutes no direct demonstration that change in behavior is least when profit is greatest, but if, whenever a man's behavior brought him a balance of reward and cost, he changed his behavior away from what got him, under the circumstances, the less profit, there might well come a time when his behavior would not change further. That is, his behavior would be stabilized, at least for the time being. And, so far as this were true for every member of a group, the group would have a social organization in equilibrium.

I do not say that a member would stabilize his behavior at the point of greatest conceivable profit to himself, because his profit is partly at the mercy of the behavior of others. It is a commonplace that the short-run pursuit of profit by several persons often lands them in positions where all are worse off than they might conceivably be. I do not say that the paths of behavioral change in which a member pursues his profit under the condition that others are pursuing theirs too are easy to describe or predict; and we can readily conceive that in jockeying for posi-

tion they might never arrive at any equilibrium at all.

Distributive Justice

Yet practical equilibrium is often observed, and thus some further condition may make its attainment, under some circumstance, more probable than would the individual pursuit of profit left to itself. I can offer evidence for this further condition only in the behavior of subgroups and not in that of individuals. Suppose that there are two subgroups, working close together in a factory, the job of one being somewhat different from that of the other. And suppose that the members of the first complain and say: "We are getting the same pay as they are. We ought to get just a couple of dollars a week more to show that our work is more responsible." When you ask them what they mean by "more responsible," they say that, if they do their work wrong, more damage can result, and so they are under more pressure to take care.¹⁶ Something like this is a common feature of industrial behavior. It is at the heart of disputes not over absolute wages but over wage differentials—indeed, at the heart of disputes over rewards other than wages.

In what kind of proposition may we express observations like these? We may say that wages and responsibility give status in the group, in the sense that a man who takes high responsibility and gets high wages is admired, other things equal. Then, if the members of one group score higher on responsibility than do the members of another, there is a felt need on the part of the first to score higher on pay too. There is a pressure, which shows itself in complaints, to bring the *status factors*, as I have called them, into line with one another. If they are in line, a condition of *status congruence* is said to exist. In this condition the workers may find their jobs dull or irksome, but they will not complain about the relative position of groups.

But there may be a more illuminating way of looking at the matter. In my example I have considered only responsibility and pay, but these may be enough, for they represent the two kinds of thing that come into the

problem. Pay is clearly a reward: responsibility may be looked on, less clearly, as a cost. It means constraint and worry—or peace of mind foregone. Then the proposition about status congruence becomes this: If the costs of the members of one group are higher than those of another, distributive justice requires that their rewards should be higher too. But the thing works both ways: If the rewards are higher, the costs should be higher too. This last is the theory of *noblesse oblige*, which we all subscribe to, though we all laugh at it, perhaps because the *noblesse* often fails to *oblige*. To put the matter in terms of profit: though the rewards and costs of two persons or the members of two groups may be different, yet the profits of the two—the excess of reward over cost—should tend to equality. And more than “should.” The less-advantaged group will at least try to attain greater equality, as, in the example I have used, the first group tried to increase its profit by increasing its pay.

I have talked of distributive justice. Clearly, this is not the only condition determining the actual distribution of rewards and costs. At the same time, never tell me that notions of justice are not a strong influence on behavior, though we sociologists often neglect them. Distributive justice may be one of the conditions of group equilibrium.

Exchange and Social Structure

I shall end by reviewing almost the only study I am aware of that begins to show in detail how a stable and differentiated social structure in a real-life group might arise out of a process of exchange between members. This is Peter Blau’s description of the behavior of sixteen agents in a federal law-enforcement agency.¹⁷

The agents had the duty of investigating firms and preparing reports on the firms’ compliance with the law. Since the reports might lead to legal action against the firms, the agents had to prepare them carefully, in the proper form, and take strict account of the many regulations that might apply. The agents were often in doubt what they should do, and then they were supposed to take the

question to their supervisor. This they were reluctant to do, for they naturally believed that thus confessing to him their inability to solve a problem would reflect on their competence, affect the official ratings he made of their work, and so hurt their chances for promotion. So agents often asked other agents for help and advice, and, though this was nominally forbidden, the supervisor usually let it pass.

Blau ascertained the ratings the supervisor made of the agents, and he also asked the agents to rate one another. The two opinions agreed closely. Fewer agents were regarded as highly competent than were regarded as of middle or low competence; competence, or the ability to solve technical problems, was a fairly scarce good. One or two of the more competent agents would not give help and advice when asked, and so received few interactions and little liking. A man that will not exchange, that will not give you what he has when you need it, will not get from you the only thing you are, in this case, able to give him in return, your regard.

But most of the more competent agents were willing to give help, and of them Blau says:

A consultation can be considered an exchange of values: both participants gain something, and both have to pay a price. The questioning agent is enabled to perform better than he could otherwise have done, without exposing his difficulties to his supervisor. By asking for advice, he implicitly pays his respect to the superior proficiency of his colleague. This acknowledgment of inferiority is the cost of receiving assistance. The consultant gains prestige, in return for which he is willing to devote some time to the consultation and permit it to disrupt his own work. The following remark of an agent illustrates this: ‘I like giving advice. It’s flattering, I suppose, if you feel that others come to you for advice.’¹⁸

Blau goes on to say: “All agents liked being consulted, but the value of any one of very many consultations became deflated for experts, and the price they paid in frequent interruptions became inflated.”¹⁹ This implies that, the more prestige an agent received, the

less was the increment of value of that prestige; the more advice an agent gave, the greater was the increment of cost of that advice, the cost lying precisely in the forgone value of time to do his own work. Blau suggests that something of the same sort was true of an agent who went to a more competent colleague for advice: the more often he went, the more costly to him, in feelings of inferiority, became any further request. "The repeated admission of his inability to solve his own problems . . . undermined the self-confidence of the worker and his standing in the group."²⁰

The result was that the less competent agents went to the more competent ones for help less often than they might have done if the costs of repeated admissions of inferiority had been less high and that, while many agents sought out the few highly competent ones, no single agent sought out the latter much. Had they done so (to look at the exchange from the other side), the costs to the highly competent in interruptions to their own work would have become exorbitant. Yet the need of the less competent for help was still not fully satisfied. Under these circumstances they tended to turn for help to agents more nearly like themselves in competence. Though the help they got was not the most valuable, it was of a kind they could themselves return on occasion. With such agents they could exchange help and liking, without the exchange becoming on either side too great a confession of inferiority.

The highly competent agents tended to enter into exchanges, that is, to interact with many others. But, in the more equal exchanges I have just spoken of, less competent agents tended to pair off as partners. That is, they interacted with a smaller number of people, but interacted often with these few. I think I could show why pair relations in these more equal exchanges would be more economical for an agent than a wider distribution of favors. But perhaps I have gone far enough. The final pattern of this social structure was one in which a small number of highly competent agents exchanged advice for prestige with a large number of others less competent and in which the less competent agents exchanged, in pairs and in

trios, both help and liking on more nearly equal terms.

Blau shows, then, that a social structure in equilibrium might be the result of a process of exchanging behavior rewarding and costly in different degrees, in which the increment of reward and cost varied with the frequency of the behavior, that is, with the frequency of interaction. Note that the behavior of the agents seems also to have satisfied my second condition of equilibrium: the more competent agents took more responsibility for the work, either their own or others', than did the less competent ones, but they also got more for it in the way of prestige. I suspect that the same kind of explanation could be given for the structure of many "informal" groups.

Summary

The current job of theory in small-group research is to make the connection between experimental and real-life studies, to consolidate the propositions that empirically hold good in the two fields, and to show how these propositions might be derived from a still more general set. One way of doing this job would be to revive and make more rigorous the oldest of theories of social behavior—social behavior as exchange.

Some of the statements of such a theory might be the following. Social behavior is an exchange of goods, material goods but also non-material ones, such as the symbols of approval or prestige. Persons that give much to others try to get much from them, and persons that get much from others are under pressure to give much to them. This process of influence tends to work out at equilibrium to a balance in the exchanges. For a person engaged in exchange, what he gives may be a cost to him, just as what he gets may be a reward, and his behavior changes less as profit, that is, reward less cost, tends to a maximum. Not only does he seek a maximum for himself, but he tries to see to it that no one in his group makes more profit than he does. The cost and the value of what he gives and of what he gets vary with the quantity of what he gives and gets. It is surprising how familiar these propositions are; it is sur-

prising, too, how propositions about the dynamics of exchange can begin to generate the static thing we call “group structure” and, in so doing, generate also some of the propositions about group structure that students of real-life groups have stated.

In our unguarded moments we sociologists find words like “reward” and “cost” slipping into what we say. Human nature will break in upon even our most elaborate theories. But we seldom let it have its way with us and follow up systematically what these words imply.²¹ Of all our many “approaches” to social behavior, the one that sees it as an economy is the most neglected, and yet it is the one we use every moment of our lives—except when we write sociology.

Endnotes

1. See R. B. Braithwaite, *Scientific Explanation* (Cambridge: Cambridge University Press, 1953).
2. Translated by I. Cunnison as *The Gift* (Glencoe, Ill.: Free Press, 1954).
3. In social anthropology D. L. Oliver is working along these lines, and I owe much to him. See also T. M. Newcomb, “The Prediction of Interpersonal Attraction,” *American Psychologist*, XI (1956), 575–86.
4. B. F. Skinner, *Science and Human Behavior* (New York: Macmillan Co., 1953).
5. *Ibid.*, pp. 297–329. The discussion of “double contingency” by T. Parsons and E. A. Shils could easily lead to a similar paradigm (see *Toward a General Theory of Action* [Cambridge, Mass.: Harvard University Press, 1951], pp. 14–16).
6. K. W. Back, “The Exertion of Influence through Social Communication,” in L. Festinger, K. Back, S. Schachter, H. H. Kelley, and J. Thibaut (eds.), *Theory and Experiment in Social Communication* (Ann Arbor: Research Center for Dynamics, University of Michigan, 1950), pp. 21–36.
7. S. Schachter, N. Ellertson, D. McBride, and D. Gregory, “An Experimental Study of Cohe-
- siveness and Productivity,” *Human Relations*, IV (1951), 229–38.
8. Skinner, *op. cit.*, p. 100.
9. S. Schachter, “Deviation, Rejection, and Communication,” *Journal of Abnormal and Social Psychology*, XLVI (1951), 190–207.
10. L. Festinger, S. Schachter, and K. Back, *Social Pressures in Informal Groups* (New York: Harper & Bros., 1950), pp. 72–100.
11. For propositions holding good of groups in practical equilibrium see G. C. Homans, *The Human Group* (New York: Harcourt, Brace & Co., 1950), and H. W. Riecken and G. C. Homans, “Psychological Aspects of Social Structure,” in G. Lindzey (ed.), *Handbook of Social Psychology* (Cambridge, Mass.: Addison-Wesley Publishing Co., 1954), II, 786–832.
12. See Homans, *op. cit.*, pp. 244–48, and R. F. Bales, “The Equilibrium Problem in Small Groups,” in A. P. Hare, E. F. Borgatta, and R. F. Bales (eds.), *Small Groups* (New York: A. A. Knopf, 1953), pp. 450–56.
13. Homans, *op. cit.*, pp. 281–301.
14. *Human Relations*, VII (1954), 313–25.
15. G. J. Stigler, *The Theory of Price* (rev. ed.; New York: Macmillan Co., 1952), p. 99.
16. G. C. Homans, “Status among Clerical Workers,” *Human Organization*, XII (1953), 5–10.
17. Peter M. Blau, *The Dynamics of Bureaucracy* (Chicago: University of Chicago Press, 1955), 99–116.
18. *Ibid.*, p. 108.
19. *Ibid.*, p. 108.
20. *Ibid.*, p. 109.
21. *The White-Collar Job* (Ann Arbor: Survey Research Center, University of Michigan, 1953), pp. 115–27.

George Homans, “Social Behavior as Exchange.” *The American Journal of Sociology*, 63:6 (1958) pp. 597–606. Copyright © 1958 by The University of Chicago. Reprinted with permission of The University of Chicago Press. ♦

Power-Dependence Relations

Richard M. Emerson

Early in his career, Richard Emerson (1925–1982) sought to advance exchange theory by developing a parsimonious theory of power/dependence that was amenable to a quantitative research program. Unlike many sociologists, he considered laboratory experiments with humans to hold considerable potential in developing such a program. In so doing, he extended Homans' efforts to ground social exchange theory in behaviorist psychology. In this classic essay—one of the more widely cited articles in sociology—Emerson attempts to provide a link among the concepts of "power," "authority," "legitimacy," and "structure" by articulating a view of power that emphasizes its relational character. More specifically, he sees power as predicated on "ties of mutual dependence." At its most fundamental level, this means that B's power is proportionally related to A's dependency on B for particular rewards or resources. Power is a potential that is realized in social exchange. Much of Emerson's interest in this essay focuses on exchanges that he calls "balancing operations." These operations consist of various options that A can attempt in responding to situations in which power is unbalanced in favor of B. Moreover, these operations constitute the core focus of the research agenda he stakes out.

Judging from the frequent occurrence of such words as *power, influence, dominance* and *submission, status* and *authority*, the im-

portance of power is widely recognized, yet considerable confusion exists concerning these concepts.¹ There is an extensive literature pertaining to power, on both theoretical and empirical levels, and in small group² as well as large community contexts.³ Unfortunately, this already large and rapidly growing body of research has not achieved the cumulative character desired. Our *integrated* knowledge of power does not significantly surpass the conceptions left by Max Weber.⁴

This suggests that there is a place at this moment for a systematic treatment of social power. The underdeveloped state of this area is further suggested by what appears, to this author, to be a recurrent flaw in common conceptions of social power; a flaw which helps to block adequate theoretical development as well as meaningful research. That flaw is the implicit treatment of power as though it were an attribute of a person or group ("X is an influential person," "Y is a powerful group," etc.). Given this conception, the natural research question becomes "Who in community X are the power holders?" The project then proceeds to rank-order persons by some criterion of power, and this ordering is called the *power-structure*. This is a highly questionable representation of a "structure," based upon a questionable assumption of generalized power.⁵

It is commonly observed that some person X dominates Y, while being subservient in relations with Z. Furthermore, these power relations are frequently intrasensitive! Hence, to say that "X has power" is vacant, unless we specify "over whom." In making these necessary qualifications we force ourselves to face up to the obvious: power is a property of the social relation; it is not an attribute of the actor.⁶

In this paper, an attempt is made to construct a sample theory of the power aspects of social relations. Attention is focused upon characteristics of the relationship as such, with little or no regard for particular features of the persons or groups engaged in such relations. Personal traits, skills or possessions (such as wealth) which might be relevant in one relation are infinitely variable across the set of possible relations, and hence have no place in a general theory.

The Power-Dependence Relation

While the theory presented here is anchored most intimately in small groups research, it is meant to apply to more complex community relations as well. In an effort to make these conceptions potentially as broadly applicable as possible, we shall speak of relations among *actors*, where an actor can be a person or a group. Unless otherwise indicated, any relation discussed might be a person-person, group-person or group-group relation.

Social relations commonly entail *ties of mutual dependence* between the parties. A depends upon B if he aspires to goals or gratifications whose achievement is facilitated by appropriate actions on B's part. By virtue of mutual dependency, it is more or less imperative to each party that he be able to control or influence the other's conduct. At the same time, these ties of mutual dependence imply that each party is in position, to some degree, to grant or deny, facilitate or hinder, the other's gratification. Thus, it would appear that the power to control or influence the other resides in control over the things he values, which may range all the way from oil resources to ego-support, depending upon the relation in question. In short, *power resides implicitly in the other's dependency*. When this is recognized, the analysis will of necessity revolve largely around the concept of dependence.⁷

Two variables appear to function jointly in fixing the dependence of one actor upon another. Since the precise nature of this joint function is an empirical question our proposition can do no more than specify the directional relationships involved:

Dependence (Dab). The dependence of actor A upon actor B is (1) directly proportional to A's *motivational investment* in goals mediated by B, and (2) inversely proportional to the *availability* of those goals outside of the A-B relation.

In this proposition "goal" is used in the broadest possible sense to refer to gratifications consciously sought as well as rewards unconsciously obtained through the relationship. The "availability" of such goals outside of the relation refers to alternative

avenues of goal-achievement, most notably other social relations. The costs associated with such alternatives must be included in any assessment of dependency.⁸

If the dependence of one party provides the basis for the power of the other, that power must be defined as a potential influence:

Power (Pab). The power of actor A over actor B is the amount of resistance on the part of B which can be potentially overcome by A.

Two points must be made clear about this definition. First, the power defined here will not be, of necessity, observable in every interactive episode between A and B, yet we suggest that it exists nonetheless as a potential, to be explored, tested, and occasionally employed by the participants. Pab will be empirically manifest only if A makes some demand, and only if this demand runs counter to B's desires (resistance to be overcome). Any operational definition must make reference to *change* in the conduct of B attributable to demands made by A.

Second, we define power as the "resistance" which can be overcome, without restricting it to any one domain of action. Thus, if A is dependent upon B for love and respect, B might then draw A into criminal activity which he would normally resist. The reader might object to this formulation, arguing that special power is in fact restricted to certain channels. If so, the reader is apparently concerned with "legitimized power" embedded in a social structure. Rather than begin at this more evolved level, we hope to derive legitimized power in the theory itself.

The premise we began with can now be stated as $P_{ab}=D_{ba}$; the power of A over B is equal to, and based upon, the dependence of B upon A.⁹ Recognizing the reciprocity of social relations, we can represent a power-dependence relation as a pair of equations:

$$P_{ab}=D_{ba}$$

$$P_{ba}=D_{ab}$$

Before proceeding further we should emphasize that these formulations have been so worded in the hope that they will apply across a wide range of social life. At a glance

our conception of dependence contains two variables remarkably like supply and demand ("availability" and "motivational investment," respectively).¹⁰ We prefer the term *dependency* over these economic terms because it facilitates broader application, for all we need to do to shift these ideas from one area of application to another is change the motivational basis of dependency. We can speak of the economic dependence of a home builder upon a loan agency as varying directly with his desire for the home, and hence capital, and inversely with the "availability" of capital from other agencies. Similarly, a child may be dependent upon another child based upon motivation toward the pleasures of collective play, the availability of alternative playmates, etc. The same generic power-dependence relation is involved in each case. The dependency side of the equation may show itself in "friendship" among playmates, in "filial love" between parent and child, in "respect for treaties" among nations. On the other side of the equation, I am sure no one doubts that mothers, lovers, children, and nations enjoy the power to influence their respective partners, within the limit set by the partner's dependence upon them.

Finally, because these concepts are meant to apply across a wide variety of social situations, operational definitions cannot be appropriately presented here. Operational definitions provide the necessary bridge between generalizing concepts on the one hand, and the concrete features of a specific research situation on the other hand. Hence, there is no one proper operational definition for a theoretical concept.¹¹

Balance and Imbalance

The notion of reciprocity in power-dependency relations raises the question of equality or inequality of power in the relation. If the power of A over B (P_{ab}) is confronted by equal opposing power of B over A, is power then neutralized or canceled out? We suggest that in such a balanced condition, power is in no way removed from the relationship. A pattern of "dominance" might not emerge in the interaction among

these actors, but that does not imply that power is inoperative in either or both directions. A *balanced* relation and an *unbalanced* relation are represented respectively as follows:

$$\begin{array}{cc} P_{ab}=D_{ba} & P_{ab}=D_{ba} \\ || \quad || & \vee \quad \vee \\ P_{ba}=D_{ab} & P_{ba}=D_{ab} \end{array}$$

Consider two social relations, both of which are balanced, but at *different levels* of dependency (say Loeb and Leopold, as compared with two casual friends). A moment's thought will reveal the utility of the argument that balance does not neutralize power, for each party, may continue to exert profound control over the other. It might even be meaningful to talk about the parties being controlled by the relation itself.

Rather than canceling out considerations of power, reciprocal power provides the basis for studying three more features of power-relations: first, a power advantage can be defined as P_{ab} minus P_{ba} , which can be either positive or negative (a power disadvantage);¹² second, the *cohesion* of a relationship can be defined as the average of D_{ab} and D_{ba} , though this definition can be refined;¹³ and finally, it opens the door to the study *balancing operations* as structural changes in power-dependence relations which tend to reduce power advantage.

Discussion of balancing tendencies should begin with a concrete illustration. In the unbalanced relation represented symbolically above, A is the more powerful party because B is the more dependent of the two. Let actor B be a rather "unpopular" girl, with puritanical upbringing, who wants desperately to date; and let A be a young man who occasionally takes her out, while dating other girls as well. (The reader can satisfy himself about A's power advantage in this illustration by referring to the formulations above.) Assume further that A "discovers" this power advantage, and, in exploring for the limits of his power, makes sexual advances. In this simplified illustration, these advances should encounter resistance in B's puritanical values. Thus, when a power advantage is used, the weaker member will

achieve one value at the expense of other values.

In this illustration the tensions involved in an unbalanced relation need not be long endured. They can be reduced in either of two ways: (1) the girl might reduce the psychic costs involved in continuing the relation by redefining her moral values, with appropriate rationalizations and shifts in reference group attachments; or (2) she might renounce the value of dating, develop career aspirations, etc., thus reducing A's power. Notice that the first solution does not of necessity alter the unbalanced relation. The weaker member has sidestepped one painful demand but she is still vulnerable to new demands. By contrast, the second solution alters the power relation itself, in general, it appears that an unbalanced relation is unstable for it encourages the use of power which in turn sets in motion processes which we will call (a) cost reduction and (b) balancing operations.¹⁴

Cost Reduction

The "cost" referred to here amounts to the "resistance" to be overcome in our definition of power—the cost involved for one party in meeting the demands made by the other. The process of cost reduction in power dependence relations shows itself in many varied forms. In the courting relation above it took the form of alteration in moral attitudes on the part of a girl who wanted to be popular; in industry it is commonly seen as the impetus for improved plant efficiency and technology in reducing the cost of production. What we call the "mark of oppression" in the character structure of members of low social castes (the submissive and "painless" loss of freedom) might well involve the same power processes, as does the "internalization of parental codes" in the socialization process. In fact, the oedipal conflict might be interpreted as a special case of the tensions of imbalance in a power-dependence relation, and cost reduction takes the form of identification and internalization classically described. "Identification with the aggressor" in any context would appear to be explainable in terms of cost reduction.

In general, *cost reduction* is a process involving change in values (personal, social, economic) which reduces the pains incurred in meeting the demands of a powerful other. It must be emphasized, however, that these adjustments do not necessarily alter the balance or imbalance of the relation, and, as a result, they must be distinguished from more fundamental *balancing operations* described below. It must be recognized that cost reducing tendencies will take place under conditions of balance, and while this is obvious in economic transactions, it is equally true of other social relations, where the "costs" involved are anchored in modifiable attitudes and values. The intense cohesion of a lasting social relation like the Loeb-Leopold relation mentioned above can be attributed in part to the cost reduction processes involved in the progressive formation of their respective personalities, taking place in the interest of preserving the valued relation. We suggest that cost reducing tendencies generally will function to deepen and stabilize social relations over and above the condition of balance.

Balancing Operations

The remainder of this paper will deal with balancing processes which operate through changes in the variables which define the structure of the power-dependence relation as such. The formal notation adopted here suggests exactly four generic types of balancing operation. In the unbalanced relation

$$\begin{array}{cc} P_{ab}=D_{ba} & P_{ab}=D_{ba} \\ || \quad || & \vee \quad \vee \\ P_{ba}=D_{ab} & P_{ba}=D_{ab} \end{array}$$

balance can be restored either by an increase in D_{ab} or by a decrease in D_{ba} . If we recall that *dependence* is a joint function of two variables, the following alterations will move the relation toward a state of balance:

1. If B reduces motivational investment in goals mediated by A;
2. If B cultivates alternative sources for gratification of those goals;

3. If A increases motivational investment in goals mediated by B;
4. If A is denied alternative sources for achieving those goals.

While these four types of balancing operation are dictated by the logic of the scheme, we suggest that each corresponds to well known social processes. The first operation yields balance through motivational withdrawal by B, the weaker member. The second involves the cultivation of alternative social relations by B. The third is based upon "giving status" to A, and the fourth involves coalition and group formation.

In some of these processes the role of power is well known, while in others it seems to have escaped notice. In discussing any one of these balancing operations it must be remembered that a prediction of which one (or what combination) of the four will take place must rest upon analysis of conditions involved in the concrete case at hand.

In the interest of simplicity and clarity, we will illustrate each of the four generic types of balancing operation in relations among children in the context of play. Consider two children equally motivated toward the pleasures of collective play and equally capable of contributing to such play. These children, A and B, form a balanced relation if we assume further that each has the other as his only playmate, and the give-and-take of their interactions might well be imagined, involving the emergence of such equalitarian rules as "taking turns," etc. Suppose now that a third child, C, moves into the neighborhood and makes the acquaintance of A, but *not* B. The A-B relation will be thrown out of balance by virtue of A's decreased dependence upon B. The reader should convince himself of this fact by referring back to the proposition on dependence. Without any of these parties necessarily "understanding" what is going on, we would predict that A would slowly come to dominate B in the pattern of their interactions. On more frequent occasions B will find himself deprived of the pleasures A can offer, thus slowly coming to sense his own dependency more acutely. By the same token A will more frequently find B saying "yes" instead of "no" to his proposals,

and he will gain increased awareness of his power over B. The growing self-images of these children will surely reflect and perpetuate this pattern.

Operation Number One: Withdrawal

We now have the powerful A making demands on the dependent B. One of the processes through which the tensions in the unbalanced A-B relation can be reduced is *motivational withdrawal* on the part of B, for this will reduce *Db* and *Pab*. In this illustration, child B might lose some of his interest in collective play under the impact of frustrations and demands imposed by A. Such a withdrawal from the play relation would presumably come about if the other three balancing operations were blocked by the circumstances peculiar to the situation. The same operation was illustrated above in the case of the girl who might renounce the value of dating. It would seem to be involved in the dampened level of aspiration associated with the "mark of oppression" referred to above.

In general, the denial of dependency involved in this balancing operation will have the effect of moving actors away from relations which are unbalanced to their disadvantage. The actor's motivational orientations and commitments toward different areas of activity will intimately reflect this process.

Operation Number Two: Extension of Power Network

Withdrawal as a balancing operation entails subjective alterations in the weaker actor. The second operation takes place through alterations in a structure we shall call a *power network*, defined as two or more connected power-dependence relations. As we have seen in our illustration, when the C-A relation is connected through A with the A-B relation, forming a simple linear network C-A-B, the properties of A-B are altered. In this example, a previously balanced A-B relation is thrown out of balance, giving A a power advantage. This points up the general

fact that while each relation in a network will involve interactions which appear to be independent of other relations in the network (e.g., A and B are seen to play together in the absence of C; C and A in the absence of B), the internal features of one relation are nonetheless a function of the entire network. Any adequate conception of a "power structure" must be based upon this fact.

In this illustration the form of the network throws both relations within it out of balance, thus stimulating one or several of the balancing operations under discussion. If balancing operation number two takes place, *the network* will be extended by the formation of new relationships. The tensions of imbalance in the A-B and A-C relations will make B and C "ready" to form new friendships (1) with additional children D and E, thus lengthening a linear network, or (2) with each other, thus "closing" the network. It is important to notice that the lengthened network balances some relations, but not the network as a whole, while the closed network is completely balanced under the limiting assumptions of this illustration. Thus, we might offer as a corollary to operation number two: Power networks tend to achieve closure.¹⁵

If the reader is dissatisfied with this illustration in children's play relations, let A be the loan agent mentioned earlier, and B, C, . . . be home builders or others dependent upon A for capital. This is the familiar monopoly situation with the imbalance commonly attributed to it. As a network, it is a set of relations connected only at A. Just as the children were "ready" to accept new friends, so the community of actors B, C, . . . is ready to receive new loan agencies. Balancing operation number 2 involves in all cases the *diffusion* of dependency into new relations in a network. A final illustration of this principle can be found in institutionalized form in some kinship systems involving the extended family. In the case of Hopi, for example, Dorothy Eggan has described at length the diffusion of child dependency among many "mothers," thus draining off much of the force of oedipal conflicts in that society.¹⁶ We have already suggested that oedipal conflict may be taken as a special case of the ten-

sion of imbalance, which in this case appears to be institutionally handled in a manner resembling operation number two. This is not to be taken, however, as an assertion that the institution evolved as a balancing process, though this is clearly open for consideration.

It is convenient at this juncture to take my balancing operation number 4, leaving number 3 to the last.

Operation Number Four: Coalition Formation

Let us continue with the same illustration. When the B-C relation forms, closing the C-A-B network in the process of balancing, we have what appears to be a coalition of the two weaker against the one stronger. This, however, is not technically the case, for A is not involved in the B-C interactions; he simply exists as an alternative playmate for both B and C.

The proper representation of coalitions in a triad would be (AB)-C, (AC)-B, or (BC)-A. That is, a triadic network reduces to a coalition only if two members unite as a single actor in the process of dealing directly with the third. The difference involved here may be very small in behavioral terms and the distinction may seem overly refined, but it goes to the heart of an important conceptual problem (the difference between a closed "network" and a "group"), and it rests upon the fact that two very different balancing operations are involved. The C-A-B network is balanced through the addition of a third relation (C-B) in operation number two, but it is still just a power network. In operation number 4 it achieves balance through collapsing the two-relational network into one group-person relation with the emergence of a "collective actor." Operation number two reduces the power of the stronger actor, while number 4 increases the power of weaker actors through collectivization. If the rewards mediated by A are such that they can be jointly enjoyed by B and C, then the tensions of imbalance in the A-B and A-C relations can be resolved in the (BC)-A coalition.

In a general way, Marx was asking for balancing operation number 4 in his call to "Workers of the world," and the collectivization of labor can be taken as an illustration of this balancing tendency as an historic process. Among the balancing operations described here, coalition formation is the one most commonly recognized as a power process. However, the more general significance of this balancing operation seems to have escaped notice, for the typical coalition is only one of the many forms this same operation takes. For this reason the next section will explore coalition processes further.

The Organized Group

We wish to suggest that the coalition process is basically involved in all organized group functioning, whether the group be called a coalition or not. We believe this illuminates the role which power processes play in the emergence and maintenance of group structure in general.

In the typical coalition pattern, (AB)-C, A and B constitute a collective actor in the sense that they act as one, presenting themselves to their common environment as a single unit. A coalition, as one *type* of group, is characterized by the fact that (a) the common environment is an actor to be controlled, and (b) its unity is historically based upon efforts to achieve that control. Now, all we need do to blend this type of group with groups in general is to *dehumanize* the environmental problem which the group collectively encounters. Thus, instead of having the control of actor C as its end, the group attempts to control C in the interest of achieving X, some "group goal." Now, if C also aspires toward X, and if C is dependent upon the group for achieving X, C might well be one of the group members—any member. Thus, in a three-member group we have three coalition structures as *intra-group* relations, each representable as ([AB]-C)-X, with A, B and C interchangeable.

The situation involved here is reminiscent of the rapidly forming and reforming coalitions in unconsolidated children's play groups. As the group consolidates, these coalitions do not drop out of the picture; they

become stabilized features of group structure, and the stabilization process is identical with "norm formation." In fact, the demands made by (AB) of C in the power process within ([AB]-C) are exactly what we normally call *group norms* and *role-prescriptions*. Such norms are properly viewed as the "voice" of a collective actor, standing in coalition against the object of its demands. This reasoning suggests an idealized conception of group structure, based upon two types of collective demands:

(1) *Role-Prescriptions*. Specifications of behavior which all group members expect (demand) of one or more but not all members.

(2) *Group Norms*. Specifications of behavior which all group members expect of all group members.

Certain actions, when performed by some member or members, need not be performed by all other members to properly facilitate group functioning. These will tend to be incorporated in role-prescriptions, which, taken together, provide a division of labor in a role structure. Roles are defined and enforced through a consolidation of power in coalition formation. Likewise with group norms. Thus, the structure of a group (its norms and prescriptions) will specify the makeup of the coalition a member would face for any group-relevant act he might perform.

This conception of group structure is idealized in the sense that it describes complete consensus among members, even to the point of group identification and internalization of collective demands (members expect things of themselves in the above definitions). Balancing operations, along with cost reduction, should move group structure toward this ideal.

Authority

It should be clear that in introducing conceptions of group *structure* we have in no way digressed from our discussion of power processes, for the emergence of these structural forms is attributed directly to operation number four, closely resembling coalitions.

tion formation. Even the most formalized role-prescription is properly viewed as the “voice” of all members standing as a coalition in making its demand of the occupant of the role. Whenever a specific member finds occasion to remind another member of his “proper” job in terms of such prescriptions, he speaks with the *authority* of the group behind him; he is “authorized” to speak for them. In this sense, every member has authority of a kind (as in civil arrest), but authority is usually used to refer to power vested in an office or role. The situation is basically the same, however, in either case. The occupant of such a role has simply been singled out and commissioned more explicitly to speak for the group in the group’s dealings with its members. That authority is *limited*, power follows from logical necessity when role-prescriptions are treated as they are here. A dean, for example, can force faculty member A to turn in his grades on time because the demand is “legitimate,” that is, supported by a coalition of all other faculty members joining with the dean in making the demand. If that dean, however, were to employ sanctions in an effort to induce that member to polish the dean’s private car, the “coalition” would immediately re-form around the faculty member, as expressed in role-prescriptions defining the boundary of “legitimate power” or authority. The dean’s authority is power contained and restricted through balancing operation number four, coalition formation.

The notion of legitimacy is important, for authority is more than balanced power; it is *directed* power which can be employed (legitimately) only in channels defined by the norms of the group. A person holding such authority is commissioned; he does not simply have the right to rule or govern—he is obliged to. Thus, authority emerges as a transformation of power in a process called “legitimation,” and that process is one special case of balancing operation number four.¹⁷

Earlier in this section we referred to the common phenomenon of rapidly forming and re-forming coalitions in children’s play groups. Our reasoning suggests that it is precisely through these coalition processes that

unifying norms emerge. These fluctuating coalitions can be taken as the prototype of organized group life wherein the tempo of coalition realignment is accelerated to the point of being a blur before our eyes. Stated more accurately, the norms and prescriptions define implicitly the membership of the coalition which would either support or oppose any member if he were to perform any action relevant to those norms.

Operation Number Three: Emergence of Status

One important feature of group structure remains to be discussed: status and status hierarchies. It is interesting that the one remaining balancing operation provided in this theory takes us naturally to the emergence of status ordering. Operation number three increases the weaker member’s power to control the formerly more powerful member through increasing the latter’s motivational investment in the relation. This is normally accomplished through giving him status recognition in one or more of its many forms, from ego-gratifications to monetary differentials. The ego-rewards, such as prestige, loom large in this process because they are highly valued by many recipients while given at low cost to the giver.

The discussion of status hierarchies forces us to consider *intra*-group relations, and how this can be done in a theory which treats the group in the singular as an actor. The answer is contained in the idealized conception of group structure outlined above. That conception implies that every intra-group relation involves at once every member of the group. Thus, in a group with members A, B, C, and D, the relations A-B, A-C, etc. do not exist. Any interactions between A and B, for example, lie outside of the social system in question unless one or both of these persons “represents” the group in his actions, as in the coalition pattern discussed at length above. The relations which do exist are (ABCD)-A, (ABCD)-B, (ABCD)-C and (ABCD)-D as a minimum, plus whatever relations of the (ABCD)-(AB) type may be involved in the peculiar structure of the group in question. Thus, in a group of N members

we have theoretical reason for dealing with N group-member relations rather

than considering all of the $\frac{N(N-1)}{2}$ possible

member-member relations. Each of these group-member relations can now be expressed in the familiar equations for a power-dependence relation:

$$P_{gm_i} = D_{m_i g}$$

$$P_{m_i g} = D_{gm_i}$$

To account for the emergence of a status hierarchy within a group of N members, we start with a set of N group-member relations of this type and consider balancing operations in these relations.

Let us imagine a five member group and proceed on three assumptions: (1) *status* involves differential valuation of members (or roles) by the group, and this valuation is equivalent to, or an expression of, D_{gm_i} ; (2) a member who is highly valued by the group is highly valued in other *similar* groups he belongs to or might freely join; and (3) all five members have the same motivational investment in the group at the outset. Assumptions 2 and 3 are empirical, and when they are true they imply that D_{gm} and D_{mg} are inversely related across the N group-member relations. This in turn implies a state of imbalance of a very precarious nature so far as group stability is concerned. The least dependent member of a group will be the first

to break from the group, and these members are precisely the most valued members. It is this situation which balancing operation number three alleviates through "giving status" to the highly valued members, thus gaining the power to keep and control those members.

These ideas are illustrated with hypothetical values in Table 47-1, with imbalance represented as power advantage (PA). Balancing operations will tend to move PA toward zero, as shown in column 6 after the highly valued members A and B have come to depend upon the group for the special rewards of status, and in column 9 after the least valued members D and E have withdrawn some of their original motivational investment in the group. The table presents three stages in status crystallization, and the process of crystallization is seen as a balancing process. The final stage (columns 7, 8, and 9) should be achieved only in groups with very low membership turnover. The middle stage might well be perpetual in groups with new members continually coming in at the lower levels. In such "open" groups, status striving should be a characteristic feature and can be taken as a direct manifestation of the tensions of imbalance. In the final stage, such strivers have either succeeded or withdrawn from the struggle.

Among the factors involved in status ordering, this theory focuses attention upon the extreme importance of the *availability factor* in dependency as a determinant of sta-

Table 47-1

Hypothetical Values Showing the Relation Between D_{gm} and D_{mg} in a Group With Five Members

Member	Before Balancing			After Operation #3			After Operation #1		
	1 Dgm	2 Dmg*	3 PAgm**	4 Dgm	5 Dmg	6 PAgm**	7 Dgm	8 Dmg	9 PAgm**
A	5	1	-4	5	5	0	5	5	0
B	4	2	-2	4	4	0	4	4	0
C	3	3	0	3	3	0	3	3	0
D	2	4	2	2	4	2	2	2	0
E	1	5	4	1	5	4	1	1	0

* Assuming that all members have the same motivational investment in the group at the outset, and that highly valued members (A and B) are valued in other groups as well.

** Power Advantage $PA_{gm} = D_{mg} - D_{gm}$.

tus position and the values employed in status ordering. In considering Dgm (the relative value or importance the group attaches to member roles), it is notably difficult to rely upon a functional explanation. Is the pitcher more highly valued than the center fielder because he is functionally more important or because good pitchers are harder to find? Is the physicist valued over the plumber because of a "more important" functional contribution to the social system, or because physicists are more difficult to replace, more costly to obtain, etc.? The latter considerations involve the availability factor. We suggest here that the *values* people use in ordering roles or persons express the dependence of the system upon those roles, and that the availability factor in dependency plays the decisive part in historically shaping those values.¹⁸

Conclusion

The theory put forth in this paper is in large part contained implicitly in the ties of mutual dependence which bind actors together in social systems. Its principal value seems to be its ability to pull together a wide variety of social events, ranging from the internalization of parental codes to society-wide movements, like the collectivization of labor, in terms of a few very simple principles. Most important, the concepts involved are subject to operational formulation. Two experiments testing certain propositions discussed above led to the following results:

1. Conformity (Pgm) varies directly with motivational investment in the group;
2. Conformity varies inversely with acceptance in alternative groups;
3. Conformity is high at both status extremes in groups with membership turnover (see column 5, Table 47-1);
4. Highly valued members of a group are strong conformers *only* if they are valued by other groups as well. (This supports the notion that special status rewards are used to hold the highly valued member who does not depend heavily upon the group, and that in

granting him such rewards power is obtained over him.);

5. Coalitions form among the weak to control the strong (balancing operation number three);
6. The greatest rewards within a coalition are given to the less dependent member of the coalition (balancing operation number three, analogous to "status giving").

Once the basic ideas in this theory have been adequately validated and refined, both theoretical and empirical work must be extended in two main directions. First, the interaction process should be studied to locate carefully the factors leading to *perceived* power and dependency in self and others, and the conditions under which power, as a potential, will be employed in action. Secondly, and, in the long run, more important, will be study of *power networks* more complex than those referred to here, leading to more adequate understanding of complex power structures. The theory presented here does no more than provide the basic underpinning to the study of complex networks. There is every reason to believe that modern mathematics, graph theory in particular,¹⁹ can be fruitfully employed in the analysis of complex networks and predicting the outcome of power plays within such networks.

Endnotes

1. See the Communications by Jay Butler and Paul Harrison on "On Power and Authority: An Exchange on Concepts," *American Sociological Review*, 25 (October, 1960), pp. 731-732. That both men can be essentially correct in the points they make yet fail to reconcile these points, strongly suggests the need for conceptual development in the domain of power relations.
2. Among many studies, see Ronald Lippitt, Norman Polansky, Fritz Redl and Sidney Rosen, "The Dynamics of Power," *Human Relations*, 5 (February, 1952), pp. 37-64.
3. Floyd Hunter, *Community Power Structure*, Chapel Hill: University of North Carolina Press, 1953.
4. Max Weber, in *The Theory of Social and Economic Organization*, New York: Oxford Uni-

- versity Press, 1947, presents what is still a classic formulation of power, authority and legitimacy. However, it is characteristic of Weber that he constructs a typology rather than an organized theory of power.
5. Hence see Raymond E. Wolfinger, "Reputation and Reality in the Study of 'Community Power,'" *American Sociological Review*, 25 (October, 1960), pp. 636-644, for a well taken critical review of Floyd Hunter's work on these very points. The notion of "generalized power" which is not restricted to specific social relations, if taken literally, is probably meaningless. Power may indeed be generalized across a finite set of relations in a power network, but this notion too requires very careful analysis. Are you dealing with some kind of halo effect (reputations if you wish), or are the range and boundary of generalized power anchored in the power structure itself? These are questions which must be asked and answered.
 6. Just as power is often treated as though it were a property of the person, so leadership, conformity, etc., are frequently referred to the personal traits of "leaders," "conformers" and so on, as if there were distinguishable types of people. In a sociological perspective such behavior should be explicitly treated as an attribute of a relation rather than a person.
 7. The relation between power and dependence is given similar emphasis in the systematic treatment by J. Thibaut and H. H. Kelley, *The Social Psychology of Groups*, New York: John Wiley and Sons, 1959.
 8. The notion of "opportunity costs" in economics is a similar idea. If an employee has alternative employment opportunities, and if these opportunities have low associated cost (travel, etc.), the employee's dependence upon his current employer is reduced.
 9. In asserting that power is based upon the dependency of the other, it might appear that we are dealing with *one* of the bases of power ("reward power") listed by John R. P. French, Jr. and Bertram Raven, "The Bases of Social Power," *Studies in Social Power*, D. Cartwright, editor, Ann Arbor, Michigan: Institute for Social Research, 1959. However, careful attention to our highly generalized conception of dependence will show that it covers most if not all of the forms of power listed in that study.
 10. Professor Alfred Kuhn, Department of Economics, University of Cincinnati, has been working on a theory for power analysis soon to be published. The scheme he develops, though very similar to the one presented here, is put together in a different way. It is anchored more tightly to economic concepts, and hence its implications lead off in different directions from those presented below.
 11. Many different operational definitions can serve one theoretical concept, and there is no reason to require that they produce intercorrelated results when applied in the same research situation. While the controversies surrounding "operationalism" have now been largely resolved, there remains some confusion on this point. See, for example, Bernice Eisman, "Some Operational Measures of Cohesiveness and Their Interrelations," *Human Relations*, 12 (May, 1959), pp. 183-189.
 12. J. Thibaut and H. H. Kelley, *op. cit.*, pp. 107-108.
 13. This definition of cohesion, based upon dependency seems to have one advantage over the definition offered by Leon Festinger, et al., *Theory and Experiment in Social Communication*, Ann Arbor: Research Center for Group Dynamics, University of Michigan Press, 1950. The Festinger definition takes into account only one of the two variables involved in dependency.
 14. The "tensions of imbalance," which are assumed to make an unbalanced relation unstable, are closely related to the idea of "distributive justice" discussed by George C. Homans, *Social Behavior: Its Elementary Forms*, New York: Harcourt, Brace and World, Inc., 1961. All of what Homans has to say around this idea could be fruitfully drawn into the present formulation.
 15. The notion of closed versus open networks as discussed here can be directly related to research dealing with communication networks, such as that reported by Harold J. Leavitt, "Some Effects of Communication Patterns on Group Performance," *Journal of Abnormal and Social Psychology*, 46 (January, 1951), pp. 38-50, in which the limiting assumptions involved in this discussion are fully met by experimental controls. In discussing those experiments in terms of the concepts in this theory we would consider each actor's dependence upon other actors for information. A formal treatment of such networks is suggested by A. Bavelas, "A Mathematical Model For Group Structure," *Applied Anthropology*, 7 (Summer, 1948), pp. 16-30.

16. Dorothy Eggan, "The General Problem of Hopi Adjustment," *American Anthropologist*, 45 (July-September, 1943), pp. 357-373.
17. The process of legitimation has sometimes been described as a tactic employed by a person aspiring to power or trying to hold his power, rather than a process through which persons are granted restricted power. For example, C. Wright Mills states: "Those in authority attempt to justify their rule over institutions by linking it, as if it were a necessary consequence, with widely believed in moral symbols, sacred emblems, legal formulae. These central conceptions may refer to God or gods, the 'vote of the majority,' 'the will of the people,' 'the aristocracy of talent or wealth,' to the 'divine right of kings,' or to the allegedly extraordinary endowments of the ruler himself. Social scientists, following Weber, call such conceptions 'legitimations,' or sometimes 'symbols of justification'" (*The Sociological Imagination*, New York: Oxford University Press, 1959, p. 36). Whether we view the process of legitimation in the context of the formation of such collective conceptions, or in the context of calling upon them to justify action, the process is fundamentally that of mobilizing collective support to oppose those who challenge power.
18. "Motivational investment" and "availability," which jointly determine dependency at any point in time, are functionally related through time. This is implied in our balancing operations. While these two variables can be readily distinguished in the case of Dmg, they are too intimately fused in Dgm to be clearly separated. The values by which a group sees a given role as "important" at time 2, evolve from felt scarcity in that role and similar roles at time 1.
19. F. Harary and R. Norman, *Graph Theory as a Mathematical Model in the Social Sciences*, Ann Arbor: Institute for Social Research, 1953. An effort to apply such a model to power relations can be found in John R. P. French, Jr., "A Formal Theory of Social Power," *The Psychological Review*, 63 (May, 1956), pp. 181-194.

Reprinted from Richard M. Emerson, "Power-Dependence Relations," *American Sociological Review*, 27(1), pp. 31-41, February, 1962. Copyright © by the American Sociological Association. ♦

Human Capital and Social Capital

James S. Coleman

In *Foundations of Social Theory* (1990), a lengthy theoretical treatise written near the end of a long and varied sociological career, James S. Coleman (1926–1995) emerged as the most important spokesperson in sociology for rational choice theory, an orientation that has had a major impact in economics and political science. As with Homans' exchange theory, the starting point for Coleman's paradigm is the individual; he endorses a conceptual orientation known as "methodological individualism." The two elementary concepts in Coleman's theory are actors and resources. In this selection from the book, two key resources—human capital and social capital—are described. The former refers to the skills and knowledge an individual possesses, while the latter refers to social relations.

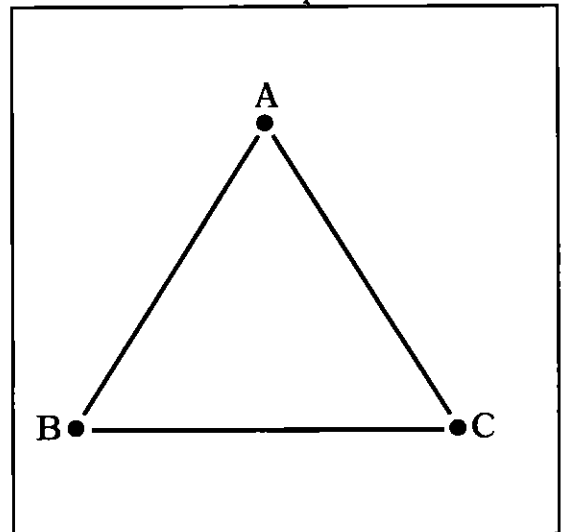
Probably the most important and most original development in the economics of education in the past thirty years has been the idea that the concept of physical capital, as embodied in tools, machines, and other productive equipment, can be extended to include human capital as well (see Schultz, 1961; Becker, 1964). Just as physical capital is created by making changes in materials so as to form tools that facilitate production, human capital is created by changing persons so as to give them skills and capabilities that make them able to act in new ways.

Social capital, in turn, is created when the relations among persons change in ways that facilitate action. Physical capital is wholly tangible, being embodied in observable material form; human capital is less tangible,

being embodied in the skills and knowledge acquired by an individual; social capital is even less tangible, for it is embodied in the *relations* among persons. Physical capital and human capital facilitate productive activity, and social capital does so as well. For example, a group whose members manifest trustworthiness and place extensive trust in one another will be able to accomplish much more than a comparable group lacking that trustworthiness and trust.

The distinction between human capital and social capital can be exhibited by a diagram such as Figure 48-1, which represents the relations of three persons (A, B, and C); the human capital resides in the nodes, and the social capital resides in the lines connecting the nodes. Social capital and human capital are often complementary. For example, if B is a child and A is an adult who is a parent of B, then for A to further the cognitive development of B, there must be capital in both the node and the link. There must be human capital held by A and social capital in the relation between A and B.

Figure 48-1
Three-Person Structure: Human Capital in Nodes and Social Capital in Relations



Using the concept of social capital will uncover no processes that are different in fundamental ways from those discussed in other chapters. This concept groups some of

those processes together and blurs distinctions between types of social relations, distinctions that are important for other purposes. The value of the concept lies primarily in the fact that it identifies certain aspects of social structure by their function, just as the concept “chair” identifies certain physical objects by their function, disregarding differences in form, appearance, and construction. The function identified by the concept “social capital” is the value of those aspects of social structure to actors, as resources that can be used by the actors to realize their interests.

By identifying this function of certain aspects of social structure, the concept of social capital aids in both accounting for different outcomes at the level of individual actors and making the micro-to-macro transition without elaborating the social-structural details through which this occurs. For example, characterizing the clandestine study circles of South Korean radical students as constituting social capital that these students can use in their revolutionary activities is an assertion that the groups constitute a resource which aids in moving the students from individual protest to organized revolt. If a resource that accomplishes this task is held to be necessary in a theory of revolt . . . then the study circles can be grouped with other organizational structures, of different origins, which have fulfilled the same function for individuals with revolutionary goals in other contexts, such as the *comités d'action lycéen* of the French student revolt of 1968 or the workers' cells in czarist Russia described and advocated by Lenin (1973 [1902]).

It is true, of course, that for other purposes one wants to investigate the details of such organizational resources, to understand the elements that are critical to their usefulness as resources for a given purpose, and to examine how they came into being in a particular case. But the concept of social capital can allow showing how such resources can be combined with other resources to produce different system-level behavior or, in other cases, different outcomes for individuals. Whether social capital will come to be as useful a quantitative concept

in social science as are the concepts of financial capital, physical capital, and human capital remains to be seen; its current value lies primarily in its usefulness for qualitative analyses of social systems and for those quantitative analyses that employ qualitative indicators.

. . . [T]he concept of social capital will be left unanalyzed (as it was in the brief descriptions given above as examples). In this chapter, however, I will examine just what it is about social relations that can constitute useful capital resources for individuals.

Obligations and Expectations

. . . [I]f A does something for B and trusts B to reciprocate in the future, this establishes an expectation in A and an obligation on the part of B to keep the trust. This obligation can be conceived of as a “credit slip” held by A to be redeemed by some performance by B. If A holds a large number of these credit slips from a number of persons with whom he has relations, then the analogy to financial capital is direct: The credit slips constitute a large body of credit on which A can draw if necessary—unless, of course, the placement of trust has been unwise, and the slips represent bad debts that will not be repaid. In some social structures (such as, for example, the neighborhoods discussed by Willmott and Young, 1967) it is said that people are “always doing things for each other.” There are a large number of these credit slips outstanding, often on both sides of a relation (for these credit slips often appear to be not fungible across different areas of activity, so credit slips from B held by A and those from A held by B are not fully used to cancel each other out). . . . In other social structures where individuals are more self-sufficient, depending on each other less, there are fewer of these credit slips outstanding at any time.

Two elements are critical to this form of social capital: the level of trustworthiness of the social environment, which means that obligations will be repaid, and the actual extent of obligations held. Social structures differ in both of these dimensions, and actors within a particular structure differ in the second.

A case which illustrates the value of trustworthiness is the rotating credit association found in Southeast Asia and elsewhere. These associations are groups of friends and neighbors who typically meet monthly; each person contributes the same amount of money to a central fund, which is then given to one of the members (through bidding or by lot). After n months each of the n persons has made n contributions and received one payout. As Geertz (1962) points out, these associations serve as efficient institutions for amassing savings for small capital expenditures, an important aid to economic development. Without a high degree of trustworthiness among the members of the group, such a credit association could not exist—for a person who received a payout early in the sequence of meetings could abscond, leaving the others with a loss. One could not imagine such a rotating credit association operating successfully in urban areas marked by a high degree of social disorganization—or, in other words, by a lack of social capital.

Another situation in which extreme trustworthiness facilitates actions that would not otherwise be possible is that of heads of state. Various accounts of the experiences of heads of state suggest that for persons in this position it is extremely valuable to have an extension of one's self, an agent one can trust absolutely to act as one would in a given situation. Many heads of state have such a person, who may not occupy a formal position of power but may be a member of a personal staff. The fact that these persons are often old friends, or cronies, rather than persons who have distinguished themselves in some political activity, is derivative from this: The most important attribute of such a person is that trust can be placed in him, and this requirement often dictates choosing a long-term personal friend. Such persons often come to have enormous power due to their proximity to a head of state and the trust placed in them; and there are many recorded accounts of the use of that power. What is of interest here is the social capital this relation provides for the head of state, assuming that the trust is well placed. The trusted other is virtually an extension of self, allowing the

head of state to expand his capacity for action.

Still another case that illustrates the importance of trustworthiness as a form of social capital is a system of mutual trust. The extreme example of such a system is a couple, each of whom places extensive trust in the other, whether they are deeply in love or not. For both members of such a couple, the relation has extraordinary psychological value. Each can confide in the other, can expose inner doubts, can be completely forthright with the other, can raise sensitive issues—all without fear of the other's misuse of the trust.

Differences in social structures with respect to the extent of outstanding obligations arise for a variety of reasons. These include, besides the general level of trustworthiness that leads obligations to be repaid, the actual needs that persons have for help, the existence of other sources of aid (such as government welfare services), the degree of affluence (which reduces the amount of aid needed from others), cultural differences in the tendency to lend aid and ask for aid (see Banfield, 1967), the degree of closure of social networks, the logistics of social contacts (see Festinger, Schachter, and Back, 1963), and other factors. Individuals in social structures with high levels of obligations outstanding at any time, whatever the source of those obligations, have greater social capital on which they can draw. The density of outstanding obligations means, in effect, that the overall usefulness of the tangible resources possessed by actors in that social structure is amplified by their availability to other actors when needed.

In a farming community such as . . . where one farmer got his hay baled by another and where farm tools are extensively borrowed and lent, the social capital allows each farmer to get his work done with less physical capital in the form of tools and equipment. Such a social structure is analogous to an industrial community in which bills of exchange (that is, debts) are passed around, serving as money and effectively reducing the financial capital necessary to carry out a given level of manufacturing activity. (See Ashton, 1945, for a description of this in

Lancashire in the 1790s, before a centralized monetary system was well established in England.)

Individual actors in a social system also differ with respect to the extent of credit slips on which they can draw at any time. For example, in hierarchically structured extended family settings, a patriarch often holds an extraordinarily large set of such credit slips, which he can call in at any time to get done what he wants done. Another clear example occurs in villages in traditional settings that are highly stratified, where certain wealthy families, because of their wealth, have built up extensive credits on which they can call at any time. (It is the existence of such asymmetries that can make some families immune to sanctions that can be used to regulate the actions of others in the community. . . .)

Similarly, in a political setting such as a legislature, a legislator in a position that brings extra resources (such as the Speaker of the House of Representatives or the Majority Leader of the Senate in the U.S. Congress) can, by effective use of those resources, build up a set of credits from other legislators so that it becomes possible for him to get legislation passed that would otherwise be defeated. This concentration of obligations constitutes social capital that is useful not only for the powerful legislator, but also in increasing the level of action of the legislature. Thus those members of legislatures who have extensive credit slips should be more powerful than those who do not because they can use the credits to produce bloc voting on many issues. It is well recognized, for example, that in the U.S. Senate, some senators are members of what is called the Senate Club, and others are not. This in effect means that some senators are embedded in a system of credits and debts, and others (outside the Club) are not. It is also well recognized that those in the Club are more powerful than those outside it.

Another example showing asymmetry in the sets of obligations and expectations is the one . . . about the crisis in medical care in the United States due to liability suits. Traditionally physicians have been in control of events having literally life-and-death impor-

tance to patients, who in turn often felt unable to adequately compensate them for the extreme benefits they brought about. Part of a physician's payment was in the form of gratitude, deference, and high occupational prestige. These constituted a felt obligation to the physician, a form of social capital which inhibited patients dissatisfied with the outcome of their medical treatments from taking action against the physician.

But several factors have changed. One is that physicians' monopoly on medical knowledge has been lessened by an expansion of education. A second is a reduction in the likelihood that there is a personal relation between physician and patient, since a patient is less likely to use a family doctor or even a general practitioner and more likely to see specialists for particular medical problems. A third is the high income of many physicians, which reduces the perceived asymmetry between service and compensation. A fourth is the increased use of liability insurance, which transfers the financial cost of a lawsuit from physician to insurer. The combination of these and other factors has reduced the social capital that protected the physician from becoming a target when patients experienced undesirable medical outcomes.

Why do rational actors create obligations? Although some of the variation in the extent of outstanding obligations arises from social changes of the sort described above, some appears to arise from the intentional creation of obligation by a person who does something for another. For example, Turnbull (1972), who studied the Ik, a poverty-ridden tribe in Africa, describes an occasion when a man arrived home to find his neighbors, unasked, on the roof of his house fixing it. Despite his not wanting this aid, he was unable to induce them to stop. In this case and others there appears to be, not the creation of obligations through necessity, but a purposive creation of obligations. The giving of gifts has been interpreted in this light (see Mauss, 1954), as have the potlatches of the Kwakiutl tribe in the Pa-

cific Northwest. In rural areas persons who do favors for others often seem to prefer that these favors not be repaid immediately, and those for whom a favor is done sometimes seem anxious to relieve themselves of the obligation.

Although the motives for freeing oneself from obligations may be readily understood (especially if the existence of obligations consumes one's attention), the motives for creating obligations toward oneself are less transparent. If there is a nonzero chance that the obligation will not be repaid, it would appear that rational persons would extend such credit only if they expect to receive something greater in return—just as a bank makes a loan only at sufficient interest to realize a profit after allowing for risk. The question then becomes whether there is anything about social obligations to make a rational person interested in establishing and maintaining such obligations on the part of others toward himself.

A possible answer is this: When I do a favor for you, this ordinarily occurs at a time when you have a need and involves no great cost to me. If I am rational and purely self-interested, I see that the importance to you of this favor is sufficiently great that you will be ready to repay me with a favor in my time of need that will benefit me more than this favor costs me—unless, of course, you are also in need at that time. This does not apply when the favor is merely the lending of money, since a unit of money holds about the same interest to a person over time.¹ When the favor involves services, expenditure of time, or some other nonfungible resource, however, or when it is of intrinsically more value to the recipient than to the donor (such as help with a task that can be done by two persons but not by one), this kind of mutually profitable exchange is quite possible. The profitability for the donor depends on the recipient's not repaying the favor until the donor is in need.

Thus creating obligations by doing favors can constitute a kind of insurance policy for which the premiums are paid in inexpensive currency and the benefit arrives as valuable currency. There may easily be a positive expected profit.

There is one more point: A rational, self-interested person may attempt to prevent others from doing favors for him or may attempt to relieve himself of an obligation at a time he chooses (that is, when repaying the favor costs him little), rather than when the donor is in need, because the call for his services may come at an inconvenient time (when repaying the obligation would be costly). Thus in principle there can be a struggle between a person wanting to do a favor for another and the other not wanting to have the favor done for him or a struggle between a person attempting to repay a favor and his creditor attempting to prevent repayment.

Information Potential

An important form of social capital is the potential for information that inheres in social relations. Information is important in providing a basis for action. But acquisition of information is costly. The minimum it requires is attention, which is always in short supply. One means by which information can be acquired is to use social relations that are maintained for other purposes. Katz and Lazarsfeld (1955) show how this operates for women in several areas of life; for example, a woman who has an interest in being in style but not at the leading edge of fashion can use certain friends, who do stay on the leading edge, as sources of information. As another example, a person who is not deeply interested in current events but who is interested in being informed about important developments can save the time required to read a newspaper if he can get the information he wants from a friend who pays attention to such matters. A social scientist who is interested in being up to date on research in related fields can make use of his everyday interactions with colleagues to do so, if he can depend on them to be up to date in their fields.

All these are examples of social relations that constitute a form of social capital in providing information that facilitates action. The relations in this case are valuable for the information they provide, not for the credit slips they provide in the form of obligations that one holds for others' performance.

Norms and Effective Sanctions

... When an effective norm does exist, it constitutes a powerful, but sometimes fragile, form of social capital. Effective norms that inhibit crime in a city make it possible for women to walk freely outside at night and for old people to leave their homes without fear. Norms in a community that support and provide effective rewards for high achievement in school greatly facilitate the school's task. A prescriptive norm that constitutes an especially important form of social capital within a collectivity is the norm that one should forgo self-interests to act in the interests of the collectivity. A norm of this sort, reinforced by social support, status, honor, and other rewards, is the social capital which builds young nations (and which dissipates as they grow older), strengthens families by leading members to act selflessly in the family's interest, facilitates the development of nascent social movements from a small group of dedicated, inward-looking, and mutually rewarding persons, and in general leads persons to work for the public good. In some of these cases the norms are internalized; in others they are largely supported through external rewards for selfless actions and disapproval for selfish actions. But whether supported by internal or external sanctions, norms of this sort are important in overcoming the public-good problem that exists in conjoint collectivities.

As all these examples suggest, effective norms can constitute a powerful form of social capital. This social capital, however, like the forms described earlier, not only facilitates certain actions but also constrains others. Strong and effective norms about young persons' behavior in a community can keep them from having a good time. Norms which make it possible for women to walk alone at night also constrain the activities of criminals (and possibly of some noncriminals as well). Even prescriptive norms that reward certain actions, such as a norm which says that a boy who is a good athlete should go out for football, are in effect directing energy away from other activities. Effective norms in an area can reduce innovativeness in that area, can constrain not only deviant actions

that harm others but also deviant actions that can benefit everyone. (See Merton, 1968, pp. 195–203, for a discussion of how this can come about.)

Authority Relations

If actor A has transferred rights of control of certain actions to another actor, B, then B has available social capital in the form of those rights of control. If a number of actors have transferred similar rights of control to B, then B has available an extensive body of social capital, which can be concentrated on certain activities. Of course, this puts extensive power in B's hands. What is not quite so straightforward is that the very concentration of these rights in a single actor increases the total social capital by overcoming (in principle, if not always entirely in fact) the free-rider problem experienced by individuals with similar interests but without a common authority. It appears, in fact, to be precisely the desire to bring into being the social capital needed to solve common problems that leads persons under certain circumstances to vest authority in a charismatic leader (as discussed . . . in Zablocki, 1980, and Scholem, 1973).

Appropriable Social Organization

Voluntary organizations are brought into being to further some purpose of those who initiate them. In a housing project built during World War II in a city in the eastern United States, there were many physical problems caused by poor construction, such as faulty plumbing, crumbling sidewalks, and other defects (Merton, n.d.). Residents organized to confront the builders and to address these problems in other ways. Later, when the problems were solved, the residents' organization remained active and constituted available social capital which improved the quality of life in the project. Residents had available to them resources that were seen as unavailable where they had lived before. (For example, despite the fact that there were fewer teenagers in the community, residents were more likely to express

satisfaction concerning the availability of babysitters.)

Members of the New York Typographical Union who were monotype operators formed a social club called the Monotype Club (Lipset, Trow, and Coleman, 1956). Later, as employers looked for monotype operators and as monotype operators looked for jobs, both found this organization to be an effective employment referral service and utilized it for this purpose. Still later, when the Progressive Party came into power in the New York Typographical Union, the Monotype Club served as an organizational resource for the ousted Independent Party. The Monotype Club subsequently served as an important source of social capital for the Independents, sustaining their party as an organized opposition while they were out of office.

In an example used earlier in this chapter, the study circles of South Korean student radicals were described as being groups of students who came from the same high school or hometown or church. In this case also, organization that was initiated for one purpose is appropriable for other purposes, constituting important social capital for the individuals who have available to them the organizational resources.

These examples illustrate the general point that organization brought into existence for one set of purposes can also aid others, thus constituting social capital that is available for use.² It may be that this form of social capital can be dissolved, with nothing left over, into elements that are discussed under other headings in this section, that is, obligations and expectations, information potential, norms, and authority relations. If so, listing this form of social capital is redundant. But the phenomenon of social organization being appropriated as existing social capital for new purposes is such a pervasive one that separate mention appears warranted.

Intentional Organization

A major use of the concept of social capital depends on its being a by-product of activities engaged in for other purposes. . . . [T]here is often little or no direct investment in social capital. There are, however, forms

of social capital which are the direct result of investment by actors who have the aim of receiving a return on their investment.

The most prominent example is a business organization created by the owners of financial capital for the purpose of earning income for them. These organizations ordinarily take the form of authority structures composed of positions connected by obligations and expectations and occupied by persons. . . . In creating such an organization, an entrepreneur or capitalist transforms financial capital into physical capital in the form of buildings and tools, social capital in the form of the organization of positions, and human capital in the form of persons occupying positions. Like the other forms of capital, social capital requires investment in the designing of the structure of obligations and expectations, responsibility and authority, and norms (or rules) and sanctions which will bring about an effectively functioning organization.

Another form of intentional organization is a voluntary association which produces a public good. For example, a group of parents whose children attend a school forms a PTA chapter where one did not exist before. This organization constitutes social capital not only for the organizers but for the school, the students, and other parents. Even if the organization serves only the original purpose for which it is organized and is not appropriated for other purposes, as is the case for organizations described in an earlier section, it serves this purpose, by its very nature, for a wider range of actors than those who initiated it. Such an organization is, concretely, of the same sort as those described earlier. The PTA is the same kind of organization as the Monotype Club, the residents' association formed to deal with faulty plumbing, and the church groups of South Korean Youth. All are voluntary associations. As it functions, however, the organization creates two kinds of by-products as social capital. One is the by-product described in the preceding section, the appropriability of the organization for other purposes. A second is the by-product described here: Because the organization produces a public good, its creation by one subset of persons makes its ben-

efits available to others as well, whether or not they participate. For example, the disciplinary standards promulgated by an active PTA change a school in ways that benefit nonparticipants as well as participants. . . .

Endnotes

1. It is interesting that, for persons whose interest in money fluctuates wildly over time, this sort of exchange is possible. In a rural county in West Virginia, the county clerk would lend money to the three town drunks when their need for money was great and then collect from them, with exorbitant interest, when they received their welfare checks, when money was of less interest to them.
2. A classic instance of this is described by Sills (1957). The March of Dimes was originally dedicated to the elimination of polio. When Salk's vaccine virtually eradicated polio, the March of Dimes organization did not go out of existence but directed its efforts toward other diseases.

References

- Ashton, T. S. 1945. The bill of exchange and private banks in Lancashire, 1790–1830. *Economic History Review* 15, nos. 1, 2:25–35.
- Banfield, E. 1967. *The moral basis of a backward society*. New York: Free Press.
- Becker, G. 1964. *Human capital*. New York: National Bureau of Economic Research, Columbia University Press.
- Festinger, L., S. Schachter, and K. Back. 1963. *Social pressures in informal groups*. Stanford: Stanford University Press.

- Geertz, C. 1962. The rotating credit association: a "middle rung" in development. *Economic Development and Cultural Change* 10:240–263.
- Katz, E., and P. F. Lazarsfeld. 1955. *Personal influence*. New York: Free Press.
- Lenin, V. I. 1973 (1902). *What is to be done?* Peking: Foreign Language Press.
- Lipset, M., M. A. Trow, and J. S. Coleman. 1956. *Union democracy*. New York: Free Press.
- Mauss, M. 1954. *The gift*. New York: Free Press.
- Merton, R. K. 1968. *Social theory and social structure*. 3rd ed. New York: Free Press.
- . n.d. Study of World War II housing projects. Unpublished manuscript. Columbia University, Department of Sociology.
- Scholem, G. 1973. *Sabbatai Sevi, the mystical messiah*. Princeton: Princeton University Press.
- Schultz, T. 1961. Investment in human capital. *American Economic Review* 51 (March):1–17.
- Sills, D. 1957. *The volunteers, means and ends in a national organization*. New York: Free Press.
- Turnbull, C. 1972. *The mountain people*. New York: Simon and Schuster.
- Willmott, P., and M. Young. 1967. *Family and class in a London suburb*. London: New English Library.
- Zablocki, B. 1980. *Alienation and charisma*. New York: Free Press.

Reprinted by permission of the publisher from *Foundations of Social Theory* by James S. Coleman, Cambridge, MA: The Belknap Press of Harvard University Press, Copyright © 1990 by the President and Fellows of Harvard College. ♦

49

The Emergence of Cooperative Social Institutions

Michael Hechter

From a rational choice perspective, Michael Hechter (b. 1943) attempts in this essay to offer an account of the manner in which social institutions arise. Rejecting what he terms the "invisible hand" approach, which treats institutions as the spontaneous outcome of the actions of self-interested individuals in interaction with others, he opts for a "solidaristic" approach. Of the two variants of solidaristic explanation—the imposition of institutions by powerful rulers versus the voluntary construction of institutions by relatively equal individuals—Hechter turns to the latter, since it raises the more interesting theoretical issues. Not the least of these issues is the matter of the "free-rider problem," a major focus of attention in this selection.

The origin of social institutions is a very old concern in social theory. Currently it has re-emerged as one of the most intensely debated issues in social science. Among economists and rational choice theorists, there is growing awareness that most, if not all, of the social outcomes that are of interest to explain are at least partly a function of institutional constraints. Yet the role of institutions is negligible both in general equilibrium theory and in most neoclassical economic models. Among other social scientists, there is a burgeoning substantive interest in institu-

tions ranging from social movements, to formal organizations, to states, and even international regimes.

This chapter discusses the two principal approaches to the problem of institutional genesis—*invisible-hand* and *solidaristic*. It further argues that the second of these is likely to afford us with a better means of attacking the problem than the first. Finally, one particular solidaristic explanation that holds promise for future research on institutional genesis is introduced.

The Concept of Social Institutions

Although the term *institution* is bandied about quite liberally in contemporary social science, no consensual definition of it has as yet emerged. The ambiguity of the term gives authors both the obligation and the license to adopt their favorite definition. At the most general level, I will take the existence of a social institution to be revealed by the appearance of some *regularity in collective behavior*. *Collective behavior* may be said to occur if different individuals behave similarly when placed in the same social situation;¹ *regularity*, for its part, indicates that this collective behavior endures over some long but indefinite period of time.

If institutions are revealed by the appearance of collective behavioral regularities, then one naturally wonders both about their origins and about the mechanisms responsible for their persistence. In institutionally rich environments, new institutions can arise from old ones through modification or diffusion processes (White, 1981; DiMaggio and Powell, 1983). Such solutions to the problem of institutional genesis are limited, however, because they are exogenous and thus beg the question of the prime mover.

What is most challenging to account for theoretically is just how institutions emerge out of anarchy, that is, from a state of nature. How, in other words, do institutions ever arise from a *noninstitutional* environment? Two types of explanations have been advanced to address this hoary old Hobbesian problem.

The *invisible-hand* approach to institutional genesis, advocated to a greater or

lesser degree by Menger [1883] (1963), Hayek (1973; 1976), and Nozick (1974), among others, views the emergence of institutions as a spontaneous by-product of the voluntary actions of self-interested individuals who share *no common ends or values* (see Hayek, 1976:111). In such accounts, existing social institutions are usually conceived as Pareto-efficient equilibria; therefore they are self-sustaining (because no one who is subject to them has an incentive to change them), rather than dependent on some third-party enforcement apparatus (like the state) whose existence, in turn, requires additional explanation.

Since invisible-hand arguments can offer an entirely endogenous explanation for the emergence of social institutions, they are to be admired for their parsimony and elegance (Nozick, 1974:18–22; Ullmann-Margalit, 1978). Their principal advantage is that they rely on fewer assumptions than do other kinds of explanations.

The alternative approach to the problem of institutional genesis rests on quite different premises. Rather than emerging spontaneously among self-interested actors each pursuing their own ends, institutions in this view are a product of *solidarity*. Solidarity can only arise among individuals who share some common end (Hechter, 1987). To attain this common end, actors must establish a set of obligations as well as a mechanism that enforces compliance to these obligations (Hobbes, [1651] 1968; Durkheim [1897] 1951; Blau, 1964:253; Hayek, 1976). From the solidaristic perspective, institutions persist not because they constitute self-enforcing equilibria, but because they are supported by consciously-designed controls.

There are two varieties of solidaristic explanations. On the one hand, institutions can be *imposed* upon a given population by some conqueror or overlord. Since it is easy to explain institutional emergence in the face of significant power differentials among individuals, this solution begs too many questions to be theoretically interesting (as Hobbes well understood). On the other hand, individuals with roughly equal power can create institutions *voluntarily*, in effect binding themselves to a joint project. This

contractarian process is theoretically interesting precisely because it is such a problematic outcome.

Which approach is superior, the invisible-hand or solidaristic one? There is a great deal of debate in the literature on this question. Most of the advocates of invisible-hand explanations of institutional genesis rest their arguments on repeated game theory.

Yet, these arguments only suffice for the establishment and maintenance of *conventions* (Lewis, 1969)—such as the rule that we all drive on the right hand side of the road²—rather than for the establishment of *n*-person *cooperative institutions*. By cooperative institution, I refer to an institution, principally serving nonclosely related kin,³ that enables those who are subject to it to reap a surplus by agreeing on a jointly maximizing strategy that is otherwise unavailable due to the absence or inappropriateness of markets.

There is an essential difference between conventions and cooperative institutions. Cooperation is the dominant strategy in conventions because there is no free-rider problem. Compliance with a convention provides its own private reward: for example, drivers who ignore conventional rules of the road take their own lives in hand. Hence, conventions indeed can be conceived of as equilibria. In cooperative institutions (which resemble Prisoner's Dilemmas), however, defection is the dominant strategy. Hence, these institutions can persist only by precluding free riders, or by assuring would-be cooperators that they are not liable to be exploited by defectors.

Contrary to the rhetoric of Taylor (1976), Hardin (1982), and Axelrod (1984), repeated game theory offers no adequate solution to the emergence of cooperation among *n* players of a Prisoner's Dilemma supergame (Hechter, 1990). The inadequacy of repeated game theory in this respect is due to two separate problems. In the first place, there are multiple equilibria in the supergame, some of which are efficient and some inefficient (Aumann, 1985).⁴ Yet under most conditions it is difficult to determine which of these multiple equilibria will be realized. In the second place, unique cooperative solutions

to the supergame rest on a most unrealistic assumption—that players are endowed with perfect monitoring capacity (Bendor and Mookherjee, 1987). This assumption limits the application of game-theoretic solutions to the evolution of cooperative institutions to the smallest of groups.⁵

In the wake of these current difficulties with the invisible-hand approach, it is best to consider the merits of solidaristic explanations, even though they require much stronger initial conditions. From a solidaristic point of view, the emergence of cooperative institutions requires individual agreement on some common end, acceptance of corporate obligations, and the establishment of formal controls to preclude free riding.

Can these admittedly strong initial conditions be explained on the basis of the typical self-interested behavioral assumptions of rational choice theory?⁶ I believe that the answer to this question is a qualified *yes*. Using the relatively weak assumptions that are traditional in rational choice, it is indeed possible to explain the emergence of cooperative institutions on the basis of solidaristic logic. The remainder of this chapter sketches out the basic argument, and then suggests that the argument can be applied to several types of empirical situations.

A Solidaristic Approach to the Emergence of Cooperative Institutions

Briefly, the genesis of cooperative institutions depends on the conjunction of (1) individuals' *demands* to provide themselves with jointly-produced private (that is, excludable) goods, as well as on (2) these individuals' potential *control capacity*—that is, their opportunities either to dissuade each other from free riding, or to assure each other of their intent to cooperate. Both demand and control capacity are necessary for the emergence of cooperative institutions; without either, this kind of institutional genesis is doomed.

The demand for cooperative institutions arises from individuals' desires to consume jointly-produced private goods (hereafter

termed *joint goods*) that cannot be obtained by following individual strategies. Cooperative institutions are generally formed to take advantage of positive externalities, such as increasing returns to scale, risk-sharing, and cost-sharing. The demand for joint goods is heightened by contextual events like wars, invasions, epidemics, and natural disasters, as well as by endogenous processes like rapid demographic growth. These events and processes are commonly experienced by a number of people, and on this account stimulate demand for goods that spread risk—such as the protection afforded by walls around a settlement, and the insurance provided the establishment of a mutual benefit society.

But the mere existence of demand for a joint good is insufficient to guarantee its production. One of the firmest conclusions of rational choice is that whereas the production of private goods is hardly problematic, in general public goods will not be produced at optimal levels, if they are produced at all. Whether a joint good is public or private is largely a function of its excludability from potential consumers. With respect to producers, both the protection afforded by town walls and the insurance offered by mutual benefit associations are *collective* goods, but with respect to consumers they are *private* goods in that these consumers (under certain conditions) can be readily excluded from them.

Whether or not a joint good is excludable is, at least in part, due to the control capacity of the potential producers of the good. This control capacity depends upon formal controls that must emerge endogenously. The establishment of these formal controls may be seen as a series of solutions to a three-tier free-rider problem. All three of these free-rider problems must be solved before a cooperative institution can emerge. Since the first two of these problems are already well-appreciated in the literature, this chapter focuses on the third of these.

The First Tier Free-Rider Problem— Design-Making

In the first place, at least one design or plan must be devised that promises to yield

the joint good. Each plan must comprise a set of *production rules* that specify what must get done by whom in order to provide an adequate supply of the good.⁷ Yet since these designs are themselves a collective good, who will devise them? Although X is eager to consume the joint good, X can spend her time more profitably by attempting to add to her resource endowment than by thinking up designs for newfangled institutions.

The solution to this first-order free-rider problem is the entrepreneurial one; it lies in the individuals' incentive to think up designs that—were their design implemented—would provide them with private benefits greatly exceeding the cost of design-making. For example, ambitious individuals would gamble by formulating plans whose adoption requires either expertise or resources that they alone can claim to have.

The Second Tier Free-Rider Problem: Establishment of an Initial Constitution

One particular design then must be selected by the relevant population. The desire to consume the joint good motivates individuals to make such a selection, for if they fail to do so, too little of the good will be produced. It is probable that each rational individual will prefer a realistic design that seems to offer the greatest amount of the good at the least (private) cost. These individual preferences must then be aggregated into a collective design. Under the conditions of the state of nature—that is, in the absence of any prior institutional framework, and in the absence of any significant resource imbalance among participants—agreement on a unanimity rule is likeliest among a relatively small group of rational egoists, because this kind of rule is most consistent with each member's private interest (Buchanan and Tullock, 1962).

The Third Tier Free-Rider Problem: Implementation of the Design

Even though all institution-builders want to consume the joint good, each rational actor will prefer to free ride on the others' contributions. This preference may not, however, characterize those contingent co-

operators who would willingly contribute to the establishment of a cooperative institution if they were assured that others would do likewise (this is often known as the *assurance problem*). If there is no means of deterring free riders, then there will be suboptimal production of the joint good—either because everyone prefers to free ride, or because the assurance problem cannot be resolved to the satisfaction of contingent cooperators.⁸

Whatever its specific causes, suboptimal production of the joint good leads the group to unravel. In order to attain optimal production, formal controls that assure high levels of compliance with production (and distribution) rules by monitoring and sanctioning group members must be adopted.

Yet since these controls are themselves a collective good, their establishment has been difficult to explain from choice-theoretic premises. One solution (the solution I have been working on) flows from the *visibility* of the production and distribution of the joint good.

For a joint good to be maximally excludable, both individual production and distribution must be highly visible. In the absence of visibility, neither free riding (a production problem), nor overconsumption (a distribution problem) can be precluded. Production visibility is at a maximum when individual effort can be well-measured by output assessment. Distribution visibility, however, is at a maximum when individuals must draw measurable shares of the joint good publicly from some central store or repository.

Most (if not all) of the positive externalities of cooperative institutions rest on the advantages of pooling individual assets so that a common central store, or bank, is thereby established. The individual depositor expects to draw some net private benefit from this central store (either interest, or—most likely in the state of nature—access to a wholly different kind of good than that deposited, such as a share of the meat of a large game animal, or insurance against some loss).

Two examples should suffice to illustrate how control is attained in cooperative insti-

tutions. In hunting and gathering societies hunters pool individual inputs of time and labor in drives to kill large game that yield meat. Both the production and distribution of killed meat is highly visible to the other hunters. The effort that each hunter contributes to the drive is difficult to conceal: individual roles in the drive are agreed upon before it takes place, and whether a given person is performing his assigned role is relatively visible (although this is a less accurate way to judge his contribution than output assessment would provide). As for distribution, the meat that is produced by the drive is usually spatially concentrated—and thereby constitutes a central fund—for, given the technology of hunter-gatherers, the most efficient way to kill large animals is to stampede them into shallow arroyos or pits (Wheat, 1967; Lee, 1979).

In rotating credit associations (Hechter, 1987:Chap. 6), individuals pool a given amount of money (which is maximally visible because it is an archetypical output) for the right to draw upon the common store of money to increase their purchasing power. In this way the monetary contributions of individual participants generate what is in effect a credit line, access to which is highly visible to all other participants.

Once individual assets are pooled in a central place, however, another free-rider problem occurs: how is it possible to stop a depositor from taking more than her fair share, or from consuming the entire central fund? This is a question that faces all rational investors—would you be likely to deposit your paycheck in a bank that you believe will soon be robbed? Presumably, only if you had some assurance that your deposit is secure. *Hence it is rational for individuals to establish formal controls in cooperative institutions so as to preserve the integrity of their investment* (which, after all, is a private good). By establishing these controls, individuals inadvertently provide themselves with a collective good—namely, security of the common fund.

But who will monitor the depositors; who will sanction them; and how will the requisite sanctioning resources be produced?

All members will take on the burden of monitoring in the initial cooperative institution. Since anyone who consumes more than their fair share of the common fund appropriates some of my own assets, I am motivated to try to get my own (augmented) investment back. There is no free-rider problem here. Whereas I can assume that other members also have an interest in getting their own investment back, I have no assurance that they won't take my share, split it among themselves, and claim that my share was never found. There is no guarantee that anyone else will look after my interests.

Likewise, *all members* will sanction the noncompliant depositor; no depositor has anything to gain by associating with a rule-breaker whose assets have already been stripped—and presumably much to lose (if it is discovered that the deviant has been helped, the helper herself is then subject to sanctioning).⁹ Finally, the ultimate sanctioning resource is easily produced, for it lies entirely within the control of the members themselves—ostracism from the group.¹⁰

By-Products of Extant Cooperative Institutions: A Fourth Tier in Institutional Genesis

It is likely that the institutionalized group may come to produce different goods than those providing its initial rationale. This is because the group now has the immense comparative advantage that it is *already organized*¹¹ and therefore can produce new joint goods much more efficiently than can unorganized individuals.¹²

In certain situations, the group may even come to produce *public* (nonexcludable) goods. This can occur if members gain so much from the production of a public good that they are willing to provide it even to non-contributors.¹³ In larger groups, this can also occur due to *agency* considerations. This will happen if the agent is not fully constrained by her principals, and if she can increase her own reputation by transforming some of the assets of the central fund into public goods.¹⁴

All told, this analysis suggests that cooperative institutions indeed can arise from the interaction of rational egoists in a state of

nature. In such an environment, however, cooperative institutions will emerge only in a contractarian fashion. Without prior cooperative institutions, there can be no entrepreneurial route to new ones. This is why the earliest institutions tend to be of the “primitive communist” variety.¹⁵ In institutionally rich environments where, for example, individual private property rights have been established, it is far simpler for these institutions to emerge via an entrepreneurial rather than a contractarian route on account of decision-making costs, and of the costs of specifying fully adequate contracts (Williamson, 1975; North, 1981).

This discussion of the emergence of cooperative institutions has two principal implications. If institutions emerge as a result of the demand for joint private goods, then *shifts in a variety of environmental and demographic conditions will heighten demand for certain kinds of joint goods and favor the emergence of institutions supplying these goods.*

Thus, the members of foraging societies tend to form local groups in the dry season—when the scarcity of water increases the benefits of cooperation among different nuclear families—but these groups disband when there is sufficient water to meet the subsistence needs of individual families (Johnson and Earle, 1987). Likewise, as markets penetrate into economically isolated territories this leads to the establishment of insurance institutions (Hechter, 1987). Finally, the rise of insecurity (due to the threat of invasion, piracy, and so forth) promotes the establishment of protective associations. Other kinds of shifts will diminish the demand for such institutions. Hence the growth of insurance markets in the late nineteenth century is associated with the decline of fraternal insurance institutions. If some public good-providing organization in a territory did not go through the first stage (that is, if it did not grow from the roots of some private good-producing institution), such evidence would contradict the thrust of this analysis.

Yet demand alone is insufficient to produce cooperative institutions: *both in their roles as producers and consumers, individuals must be highly visible to one another in*

order to reduce the severity of the free-rider and assurance problems.

In the state of nature, bulky goods that must be cooperatively acquired are likely to promote both kinds of visibility. This is consistent with the finding that meat (at least some of which is often cooperatively acquired) is more widely shared among hunter-gatherers than other types of food (Kaplan and Hill, 1985). Irrigation systems provide a graphic example of a cooperative institution that develops to provide access to a bulky joint good. Wittfogel (1957:18), for example, notes that water is a distinctive resource in that it has a tendency to gather in bulk.¹⁶ Further research into the visibility of the production of different kinds of joint goods, and of the potential centrality of these goods, doubtless will provide a richer body of empirical implications for the genesis of cooperative institutions.

It should be emphasized that the analysis in this chapter is quite different from Mancur Olson’s (1965) well-known explanation of the development of collective goods-seeking organizations like trade unions and farm organizations. Insofar as these groups sought to raise the wages of whole classes of workers, they aimed to produce a collective good. Given this, the optimal strategy for any given worker is to free ride and cash in on the (presumably successful) efforts of union organizers and their credulous followers. How, then, did these groups emerge? Olson’s explanation is that the early trade unions (in the days before the passage of closed-shop legislation) could lure members only if they provided them with desirable selective incentives, including insurance. In Olson’s account, therefore, insurance is considered to be the by-product of trade unions.

The problem with Olson’s explanation is that, like formal controls, selective incentives are themselves a collective good. This means that they, too, have to be produced by rational egoists. How is it that a group aiming to provide a public good can attract any rational members at all, let alone manage to produce selective incentives? As the previous analysis shows, the rise of groups providing immanent joint goods entails no such liability. Since they are formed for the provi-

sion of *private* goods, there is no initial free-rider problem. To obtain their goods, members are led to adopt formal controls that make possible production of the goods. Once the goods have been produced, they can be used in a variety of ways. For example, there is no inherent reason why the members of an insurance group cannot convert their common assets into a strike fund and reconstitute themselves as a trade union. . . .

Endnotes

1. Consider an elevator having male and female passengers. If male passengers are observed to allow the females to exit first, this is a collective behavioral regularity. Whether this behavioral regularity is due to the presence of a norm or to explicit rules is beside the point.
2. For the purposes of this chapter, language itself may be considered to be a convention.
3. This restriction is due to the fact that the rise of institutions among close relations can be explained easily by evolutionary arguments based on genetic relatedness. Such reasoning is, however, generally insufficient to account for institutions whose scope surpasses the members of a nuclear family.
4. The problem of multiple equilibria is double-barrelled. On the one hand, cooperation may not emerge because some of these equilibria are inefficient. On the other hand, cooperation may not emerge even if the various equilibria are all efficient, since they are unlikely to be equally preferred by all the players or the game. This situation then leads to a noncooperative bargaining problem.
5. In the absence of perfect monitoring capacity, a player can never be certain of the moves that other players have taken in past plays of the game. Thus, she cannot infer that cooperation is ever rational.
6. This question is critical, for if we suspend self-interested behavioral assumptions—and allow individuals to have internalized values or some small but positive amount of altruism—then there is an all too easy way to overcome the assurance problem, and thereby to account for the emergence of cooperative institutions. This strategy is akin to invoking a *deus ex machina*, but there can be no theoretical justification for so doing.
7. I ignore the obvious complication that the initial production functions for the joint good will be estimates, and that disagreements may well result about the accuracy of these estimates.
8. There is a growing experimental literature on the use of provision points and money-back guarantees as means of resolving the assurance problem. Whereas there is evidence that some of these arrangements do, in fact, result in the production of greater public goods, each of them is imposed *exogenously* in the experiments. Hence these solutions to the free-rider and assurance problems are inconsistent with the premises of this analysis.
9. In more complex situations where there are alternative benefit-providing institutions, deviant actors often gain a negative reputation that makes them unsuitable for admission to any such institution. After other participants get their investment back, what is their incentive to ruin the deviant's reputation? Why should rational egoists be concerned about the fortunes of the participants in *other* institutions? This kind of problem is endemic in academic hiring situations, where the members of sending departments often provide misleading information to receiving departments in hopes of getting rid of a troublesome colleague or a sub-par student. The only force that can counter this free-rider problem is the damage that such deceit might bring to the information provider in further repeat dealings. Hence, the less frequent the contact between the members of two academic departments, the less reliable the information supplied about potential colleagues and students, *ceteris paribus*. The multiplexity of ties between groups increases the probability of this intergroup sanctioning.
10. It should be noted that this solution to the emergence of cooperative institutions is practicable only in relatively small groups. In essence, the creation of a central store of resources commits participants to involvement in a repeated game. As such, many of the mechanisms that produce cooperation in the literature on repeated games (Taylor, 1976; Axelrod, 1984) are employed here to the same effect. The reader may wonder wherein this approach differs from the invisible-hand approach. Whereas repeated game theorists take the existence of the supergame (and sometimes the existence of a specific discount rate) as a given, this analysis explains how it is that rational egoists voluntarily commit themselves to social situations involving repeated exchange.

11. The connection between pre-existing organizations founded to produce joint private goods and public good-providing organizations often has been stressed in the literature on social movements (Oberschall, 1973; McAdam *et al.*, 1988). Thus in her analysis of the emergence of the contemporary women's movement, Evans (1980) locates its roots in informal networks of women who had come to know one another in the context of prior civil rights and New Left political organizations. Black churches (which offered insurance benefits) played an important crystallizing role in the development of the civil rights movement (Oberschall, 1973:126–27; McAdam, 1982; Morris, 1984). Fraternal-service groups played a similar role in the emergence of local anti-pornography movements (Curtis and Zurcher, 1973:56); and mosques played this kind of role in the early days of the Iranian Revolution (Snow and Marshall, 1984).
12. Naturally, this kind of an argument has its limits, otherwise all production would be concentrated in just one institution. For an interesting discussion of the limits of integration in firms, see Hart (1987).
13. Thus, to satisfy his desire to watch movies in the middle of the night, Howard Hughes bought a local Las Vegas station (Hardin, 1982).
14. For example, the agents of some American ethnically-based fraternal societies had political aspirations in their communities, and by judiciously investing these funds they could further these political aspirations (Stolarik, 1980). Likewise, the managers of large Minneapolis corporations are motivated to provide charitable donations in the community by the access to high prestige social circles that these donations uniquely provide (Galaskiewicz, 1985). The provision of public goods also can be a by-product of relatively homogeneous groups. In such groups, access to the joint good may be limited only to those members who contribute to specific public goods that are unrelated to the group's initial rationale. Thus some Pittsburgh fraternal associations expelled members who had committed crimes or treason, or who hired out as strikebreakers (Galey, 1977). This then explains how the self-interest of rational egoists can lead them to produce collective (and sometimes even public) goods.
15. In contrast, the Marxian explanation for primitive communism rests on questionable

arguments about the absence of a surplus beyond that necessary for subsistence.

16. Clearly, the demand for a predictable water supply is insufficient to account for actual irrigation works, for many peoples who would have gained from it did not adopt such practices. Whether the adopters of irrigation had a visibility advantage over nonadopters remains to be explored in further research.

References

- Aumann, Robert J. (1985). "Repeated Games." Pp. 209–242 in George R. Feiwel, ed., *Issues in Contemporary Microeconomics and Welfare*. Albany, NY: State University of New York Press.
- Axelrod, Robert. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- Bendor, Jonathan, and Dilip Mookherjee. (1987). "Institutional structure and the logic of ongoing collective action." *American Political Science Review*, 81(1):129–54.
- Blau, Peter M. (1964). *Exchange and Power in Social Life*. New York: John Wiley.
- Buchanan, James M. and Gordon Tullock. (1962). *The Calculus of Consent*. Ann Arbor: University of Michigan Press.
- Curtis, Russell L. and Louis A. Zurcher, Jr. (1973). "Stable resources of protest movement: The multi-organizational field." *Social Forces*, 52(1):53–60.
- DiMaggio, Paul J. and Walter W. Powell. (1983). "Institutional isomorphism." *American Sociological Review*, 48(2):147–160.
- Durkheim, Emile. (1951). *Suicide*. New York: Free Press. [Originally published 1897.]
- Evans, Sarah. (1980). *Personal Politics*. New York: Vintage Books.
- Galaskiewicz, Joseph. (1985). *Social Organization of an Urban Grants Economy: A Study of Business Philanthropy and Nonprofit Organizations*. Orlando, FL: Academic Press.
- Galey, Margaret E. (1977). "Ethnicity, fraternalism, social and mental health." *Ethnicity*, 4(1):19–53.
- Hardin, Russell. (1982). *Collective Action*. Baltimore: Johns Hopkins University Press (for Resources for the Future).
- Hart, Oliver. (1987). "Incomplete contracts and the theory of the firm." Paper presented at the Conference on Knowledge and Institutional Change, University of Minnesota.
- Hayek, Friedrich A. (1973). *Law, Legislation and Liberty*, Vol. I. Chicago: University of Chicago Press.

- (1976). *Law, Legislation and Liberty*, Vol. II. Chicago: University of Chicago Press.
- Hechter, Michael. (1987). *Principles of Group Solidarity*. Berkeley and London: University of California Press.
- (1990). "On the inadequacy of game theory for the resolution of real-world collective action problems." In M. Levi and K. S. Cook, eds., *The Limits of Rationality*. Chicago: University of Chicago Press. In press.
- Hobbes, Thomas. (1968). *Leviathan*, edited by C. B. Macpherson. Harmondsworth, England: Penguin. [Originally published 1651.]
- Johnson, Allen W. and Timothy Earle. (1987). *The Evolution of Human Societies*. Stanford: Stanford University Press.
- Kaplan, Hillard and Kim Hill. (1985). "Food sharing among Ache foragers: Tests of explanatory hypotheses." *Current Anthropology*, 26(2):223-239.
- Lee, Richard B. (1979). *The !Kung San*. Cambridge: Cambridge University Press.
- Lewis, David. (1969). *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- McAdam, Doug. (1982). *Political Process and the Development of Black Insurgency, 1930-1970*. Chicago: University of Chicago Press.
- McAdam, Doug, John D. McCarthy, and Mayer N. Zald. (1988). "Social movements and collective behavior: Building micro-macro bridges." Pp. 695-738 in Neil J. Smelser, ed., *Handbook of Sociology*. Newbury Park, CA: Sage.
- Menger, Carl. (1963). *Investigations into the Method of the Social Sciences with Special Reference to Economics*, edited by Louis Schneider. Translated by Francis J. Nock. Champaign-Urbana: University of Illinois Press. [Originally published 1883.]
- Morris, Aldon. (1984). *The Origins of the Civil Rights Movement*. New York: Free Press.
- North, Douglass C. (1981). *Structure and Change in Economic History*. New York: Norton.
- Nozick, Robert. (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Oberschall, Anthony. (1973). *Social Conflict and Social Movements*. Englewood Cliffs, N.J.: Prentice-Hall.
- Olson, Mancur. (1965). *The Logic of Collective Action*. Cambridge: Harvard University Press.
- Snow, David A. and Susan Marshall. (1984). "Cultural imperialism, social movements, and the Islamic revival." Pp. 131-152 in Louis Kriesberg, ed., *Social Movements, Conflicts and Change*, Vol. 7. Greenwich, CT: JAI Press.
- Stolarik, M. Mark. (1980). "Slovak fraternal benefit societies." Pp. 130-145 in Scott Cummings, ed., *Urban Self-Help in America*. Port Washington, NY: Kennikat Press.
- Taylor, Michael. (1976). *Anarchy and Cooperation*. London: John Wiley.
- Ullman-Margalit, Edna. (1978). "Invisible hand explanations." *Synthese*, 39:263-91.
- Wheat, J. B. (1967). "A Paleo-Indian bison kill." *Scientific American*, January.
- White, Harrison C. (1981). "Where do markets come from?" *American Journal of Sociology*, 87(3):517-547.
- Williamson, Oliver. (1975). *Markets and Hierarchies*. New York: Free Press.
- Wittfogel, Karl A. (1957). *Oriental Despotism: A Comparative Study of Total Power*. New Haven: Yale University Press.

Reprinted with permission from Michael Hechter, Karl-Dieter Opp, and Reinhard Wippler. *Social Institutions: Their Emergence, Maintenance, and Effects*. (New York: Aldine de Gruyter) Copyright © 1990 Walter de Gruyter, Inc., New York. ♦

Formulation of Exchange Theory

Peter Blau

Peter Blau (1918–2002), who was born in Vienna and emigrated to the United States during the Nazi era, was a key exponent of exchange theory for over three decades. During this time he has attempted to go beyond the general propositional stage articulated by George Homans in order to focus on social structure. In this excerpt from Structural Contexts of Opportunities (1994), Blau builds on micro-level exchange theory while articulating an appreciation of both the difference between economic and social exchange and the factors that make the macro level different from the micro level. One of the issues he addresses is the paradoxical fact that social exchange both facilitates social bonding and gives rise to status differentiation.

A fundamental difference between social life in small isolated communities and that in large complex societies is the declining significance of the groups into which one is born and the growing significance of reciprocated choices for human relations. To be sure, the significance of ascribed positions has by no means disappeared in contemporary complex societies. Most people's closest relations are with their parents and children. Other ascribed positions continue to exert a major influence on social relations, notably one's kin and the ethnic group and social class into which one is born. Yet, even for quite close relatives, except one's immediate family, the extent of social interaction and the intimacy of the relation are not ascribed but depend on reciprocal choices. Larger as-

cribed affiliations, like ethnic and class background, affect the likelihood of choice but do not predetermine who selects whom as close associate, which depends on reciprocated choices.

Thus, ascribed as well as achieved positions govern probabilities of association, which are generally higher for ascribed than achieved affiliations, but they do not determine specific associates (with the exception of parents and children), let alone the extent of social interaction and the closeness of the relation. Their probabilistic influences on ingroup associations are similar to those of a community's population structure. The population distributions in a community also influence only the probabilities of ingroup and intergroup relations of various kinds, but the specific dyads within which these probabilities find expression depend on mutual choices.

Dependence on reciprocated choice implies that, if I want to associate with someone, I cannot realize my goal unless I make him interested in associating with me. For our social relation to persist once it has been established, both of us have to sustain an interest in its continuation. To determine what brings these conditions about is the objective of exchange theory, which analyzes the processes that establish reciprocity in social relations and sustain it, and which thereby dissects the dynamics of social interaction.

Structural conditions impose limits on the exchange relations that can develop. The population structure of an entire society or large community, however, is far removed from the daily social life of individuals and hence does not affect it directly but indirectly. Multilevel structural analysis traces these indirect limiting influences. It discloses how macro-structural conditions are transmitted to successive levels and which ones reach the lowest level on which direct social interaction and exchange occur. It may indicate, for example, that society's racial heterogeneity penetrates into small substructures or that it is reflected in segregation of different races in different suburbs and neighborhoods with much homogeneity within them. The former situation would make intergroup relations more likely than

the latter, but neither would determine which specific social relations occur.

Many, if not most, human gratifications are obtained in relations with other human beings. Intellectual stimulation and relaxing conversation, sexual pleasures and the enchantment of love, academic recognition and a happy family life, satisfying the lust for power and the need for acceptance—all of these are contingent on eliciting responses from others. Exchange theory analyzes the mutual gratifications persons provide one another that sustain social relations.

The basic assumption of the theory of social exchange is that persons establish social associations because they expect them to be rewarding and that they continue social interaction and expand it because they experience it to be rewarding. This assumption that two parties associate with one another not owing to normative requirements but because they both expect rewards from doing so implies that the exchange of rewards is a starting mechanism of social relations that is not contingent on norms prescribing obligations. If a person is attracted to others because she expects associating with them to be rewarding, she will want to associate with them to obtain the expected rewards. For them to associate with her, they must be interested in doing so, which depends, according to the initial assumption, on their expecting such association to be rewarding to them. Consequently, for the first person to realize the rewards expected from the association with others, she must impress them as a desirable associate with whom interaction will be rewarding.

Individuals are often hesitant to take the first step for fear of rejection. A widely used early strategy is for people to impress others in whom they are interested with their outstanding qualities—their wit, charm, intelligence, knowledge of the arts—which implicitly promises that associating with them would be a rewarding experience. If the early steps are successful, they tend to become self-fulfilling prophecies. As each person puts his best foot forward, associating with

him turns out to be an enjoyable experience. In due course, people start doing favors for one another. In a work situation, the more experienced may give their colleagues advice or help with a difficult job. Neighbors may lend one another tools. People who met socially may issue invitations to dinner or a party.

Most people enjoy doing favors for others, usually without any thought of return, at least initially. Nevertheless, a person who benefits from an association is under an obligation to reciprocate. If the benefits are recurrent—whether involving merely the enjoyment of the other's company or getting frequently needed advice about one's work from a colleague—the self-imposed obligation to reciprocate is sustained by the interest in continuing to obtain the benefits. It is further reinforced by the fear of not seeming ungrateful. Even when there is no initial thought of return, failure to reciprocate when the occasion arises invites such an accusation, which will be experienced though it remains unspoken.

Imagine a neighbor lends you her lawn mower in the summer, but when she asks you next winter to borrow your snow blower you refuse. The neighbor and others who learn of your refusal undoubtedly will consider you ungrateful, and whether they do or not, you yourself will feel ungrateful and surely will be hesitant to ask to borrow her lawn mower again. The feelings and possible accusations of ingratitude indicate that favors freely given are not entirely free but create obligations in one's own mind to reciprocate as well as possible social pressures to discharge the obligations.

A fundamental distinction between social and economic exchange is that social exchange engenders diffuse obligations, whereas those in economic exchange are specified in an implicit or explicit contract. For economic transactions that are not immediately completed, like purchases in stores, the terms of the exchange are agreed upon in advance by both parties, and major agreements are formalized in a contract that specifies the precise nature of the obligations of both parties and when any outstanding debts are due. The favors in social

exchange, by contrast, create diffuse obligations, to be discharged at some unspecified future date. If a couple give a dinner party, for instance, they have no agreement on when and where or even whether the guests will invite them back, though their relations may be weakened if they do not, or if they do so too late or too soon. The diffuseness of the obligations implies that large-scale social exchange is not likely to occur unless firm social bonds rooted in trust have been established.

In the absence of legal obligations to make a return for benefits received, the initial problem of new acquaintances is to prove themselves trustworthy in social exchange. This typically occurs as exchange relations evolve in a slow process, starting with minor transactions entailing little risk and requiring little trust. The mutual discharge of obligations and reciprocation profit both parties and prove them increasingly trustworthy as favors are regularly reciprocated. The growing mutual advantages gained from the association fortify their social bond. This may appear to be merely a by-product of social exchange, but it is, in fact, its most important product.

Implicit in discussions of social exchange is an element of rationality, if not calculation, which may give the impression that social exchange theory is simply a version of rational choice theory. However, this impression is misleading. To be sure, social exchange does imply some rational pursuit of rewards, but the prime benefit sought, once the friendship bond of mutual support and trust is clearly established, is the rewarding experience derived from the association itself. Any material benefits exchanged are incidental and of significance largely as tokens of the friendship.

I conceptualize processes of social association as occurring in the relation between two persons. Accordingly, the exchange theory just presented analyzes exchange processes in dyads. . . . Ekeh (1974) has criticized my and Homans's (1961) exchange theory as individualistic, ignoring the differ-

ence between my concern with social structure and Homans's psychological reductionism. His criticism centers on the analysis of dyadic exchange. He contrasts the concept of restricted or two-party exchange unfavorably with Lévi-Strauss's (1949) generalized or multiparty exchange. Ekeh (1974: 62–65) considers the latter (multiparty) exchange more Durkheimian, owing to its concern with structural integration, whereas he dismisses dyadic exchange as individualistic and thus lacking a structural focus.

There is good reason that I, as a structural sociologist, prefer restricted dyadic to generalized multiparty exchange. Generalized exchange refers to the prevailing practice that all members of a tribe or group freely provide benefits to other members without looking for any return from the person to whom the contribution is made. Since doing favors for others is socially expected, it is in effect a group norm. Conformity with this norm is the reason that all group members receive favors in the long run and solidarity is strengthened. My criticism of generalized exchange is that it is simply another name for conformity to group norms and consequently commits the tautological fallacy of explaining social conduct in terms of social norms demanding this conduct.¹ Generalized exchange thereby dispenses with the crucial insight of exchange theory that interpersonal relations are not contingent on social norms, because gradually expanding reciprocity supplies a mechanism for establishing and maintaining them and engendering trust to boot.

That my analysis of social exchange is confined to exchange processes that occur in dyads does not mean that the social context in which these processes occur can be ignored, since it does influence them. Actually, exchange processes are affected by several contexts of widening social circles. The most immediate social context is the groups to which the dyads belong, which exert two distinct influences on dyadic exchange.

First, a group's network structure defines the alternative opportunities for exchange relations various persons have and thereby affects the outcomes of persons in different

network positions. (Exchange processes, in turn, may alter the network structure.) Experiments performed by Cook and her colleagues indicate that networks that provide alternative exchange partners to one person but not to others increase the bargaining power of this person in dyadic exchanges (see, for example, Cook, Gillmore, and Tamagishi 1983).

A second influence of the immediate social context is that it discourages failure to reciprocate for benefits received by social disapproval of such ingratitude. I realize that my reference to social disapproval, which implies social pressure, sounds as if I attributed exchange to group norms, for which I criticized the principle of generalized exchange. There is a major difference, however. If the practice of making a contribution freely to any group member without expecting a return from that member is explained by the cultural norm to do so, the *explicans* cannot explain the *explicandum*, because the two are redundant. But exchange is explained not by social pressures but by the returns it brings, including pleasant company or friendship as well as possibly tangible benefits. Social exchange, however, cannot prevail if trust, once established, is violated, and social disapproval discourages its violation. Social pressures do not explain—account for—reciprocal exchange, but they help to sustain it.

The influence of the wider social circles—the population structure of a neighborhood, community, or entire society—depends on the extent to which the population distributions of the encompassing social structure penetrate into the substructures of face-to-face groups. Many of the differences in society's population structure are the result of differences among rather than within substructures on successive levels. As a result, face-to-face groups are less differentiated than their encompassing social structures. Multilevel structural analysis discloses how much differentiation in various dimensions penetrates into the substructures of interpersonal relations. Greater homophily in segregated substructures promotes ingroup relations, but despite much segregation, some differentiation penetrates to the lowest

level of interpersonal relations. Consequently, although ingroup relations prevail in daily social intercourse, intergroup relations also regularly occur.

The common occurrence of intergroup relations is revealed in a study by Marsden (1990) that applies my theoretical scheme to the egocentric face-to-face networks of a sample of the American population. He initially distinguishes a demand-side view of networks in terms of preferences for various kinds of associates from a supply-side view, like my theory's, in terms of opportunities for associating with diverse others. On the basis of previous research on the composition of families and work places, we know that families are more diverse in age and sex but less diverse in ethnic and religious affiliation than associates at work. Accordingly, Marsden hypothesizes more intergroup relations in respect to age and sex and fewer intergroup relations in respect to ethnic and religious affiliation between relatives than between fellow workers. The results support these hypotheses, which stipulate intergroup as well as ingroup relations even between close associates. Marsden concludes that my macro-structural opportunity theory is applicable to the study of the relations in microstructures, contrary to what I myself had stated.

I am pleased that the theory can be used in the investigation of face-to-face networks, which I had questioned. One should note, however, that confining network analysis to the supply-side approach would fail to take full advantage of the possibilities for analysis the small scope of these networks provides. In the study of large populations, analysis and research cannot proceed without ignoring the complexities of social life by having to aggregate specific observations into gross concepts and measures, like heterogeneity, intersection, or intergroup relations. The subtle processes that govern face-to-face relations are admittedly (but inevitably) obscured by such aggregations. The study of interpersonal relations and small networks can directly analyze these processes and thereby contribute to our understanding of them.

Imbalance in Exchange

A paradox of social exchange is that it gives rise to both social bonds between peers and differentiation of status. This was the case for the ceremonial exchange of gifts in nonliterate societies, and it is the case for exchange processes in advanced industrial societies. To start by exemplifying the former, the Kula ceremonial gift exchange of the Trobriand Islanders, as discussed by Malinowski (1961: 92), "provides every man within its ring with a few friends near at hand, and with some friendly allies in far away, dangerous, foreign districts." A few pages later he states that "among the natives of the Kula . . . wealth is the indispensable appanage of social rank" (p. 97). Probably the extreme case of the significance of social exchange for differentiation of status is the famous potlatch of the Kwakiutl, a feast of reckless spending in which "status in associations and clans, and rank of every kind, are determined by the war of property" (Mauss 1954: 35).

A contemporary case of status differentiation resulting from social exchange was observed in the office of a federal government agency responsible for the enforcement of certain laws. The duties of the agents involved investigating private firms by auditing their books and conducting interviews, determining any legal violations and the action to be taken, and negotiating a settlement with the employer or a top manager. The work was quite complex, and agents often encountered problems. When they did, they were expected to consult their supervisor, but they tended to be reluctant to do so for fear of adversely affecting their annual rating by their supervisor. Instead, they usually consulted colleagues. Whereas officially prohibited, this practice was widespread and evidently tolerated. Although agents worked on different cases, one could observe all day long pairs or small clusters of persons in deep discussions, most of which dealt with problems of their cases. Lunch periods were filled with such discussions.

The observation of these consultations originally gave me the idea of social ex-

change. To cite the central passage (Blau 1955: 108):²

A consultation can be considered an exchange of values; both participants gain something and both have to pay a price. The questioning agent is enabled to perform better than he could otherwise have done, without exposing his difficulties to the supervisor. By asking for advice, he implicitly pays his respect to the superior proficiency of his colleague. This acknowledgment of inferiority is the cost of receiving assistance. The consultant gains prestige, in return for which he is willing to devote some time to the consultation and permit it to disrupt his own work. The following remark illustrates this: 'I like giving advice. It's flattering, I suppose, if you feel that the others come to you for advice.'

The principle of marginal utility applies to these exchanges. Although most agents liked being consulted, for those frequently asked for advice the gain in informal status of an additional consultation diminished and the cost in repeated interruptions of one's own work increased. As the most popular consultant said to me when asked about being consulted, "I never object, although sometimes it's annoying." The principle also applies to agents who frequently need advice, but in reverse, of course.

Repeated admissions of needing advice undermine one's self-confidence and standing in the group, particularly if an oft-interrupted consultant expresses some impatience or annoyance. To forestall such experiences, most agents establish partnerships of mutual consultation, reserving consulting the most expert colleagues for their most difficult problems. Since agents often have tentative solutions for their problems and need not so much an answer as assurance that theirs is correct, a colleague whose expertise is not superior to one's own can provide such support.

The most expert agents face a different dilemma: asking for advice or even for confirmation of their tentative solutions may well endanger their superior standing as experts. Making official decisions in a difficult case on one's own can easily raise doubts and

questions in a person's mind, even an expert's. One way to cope with this situation is to stop going over it again and again in one's head and instead telling some colleagues about the interesting problems that have arisen in a given case and discussing how they might be solved, possibly over lunch if not in the office.

Such "thinking out loud" may well stimulate new associations and ideas one would not have come up with on one's own, particularly as the listeners are also experienced agents, who might raise objections if one is on the wrong track, and whose assent implicit in attentive listening and interested questions conveys approval. In contrast to asking for advice, telling colleagues about interesting problems in a case and how they might be solved enhances the respect of one's colleagues, though it is, in effect, a subtle form of asking colleagues to corroborate one's own provisional decisions.

To put the underlying principles of imbalanced and balanced exchange into general terms, rendering important services or providing valued benefactions is a claim to superior status. Reciprocation denies this claim, and excessive returns make a counterclaim, which can lead to a potlatch-like war of seeking to outdo one another to stay ahead. Failure to reciprocate by discharging one's obligations validates the claim and acknowledges the other's superiority in return for the benefits received and in the hope of continuing to receive them. Thus, the contingency that determines whether social exchanges lead to friendships between peers or superordination and subordination is whether benefits received are reciprocated or not. This, in turn, depends on whether one of the two parties has superior resources of the kind that are in contention (which was professional competence in the case of agents).³

In a seminal article, Emerson (1962) specified conditions in which balance in social exchange can be restored. I have slightly modified his scheme to conceptualize it as four alternatives to becoming dependent on

a person's influence who has some services to offer that others need or want. First, they can give him something he needs or wants enough to reciprocate by satisfying their wishes, provided that they have resources that meet his needs. Second, they can obtain the needed benefits elsewhere, assuming that they have access to alternative sources of these benefits. These two possibilities, if recurring, result in reciprocal exchange relations between peers. Third, they can coerce him to give them what they want. This involves domination by force and is outside the purview of exchange. Fourth, they can resign themselves to do without what they thought they needed, which is Diogenes' solution for remaining independent.

If none of these four alternatives is available, the others become dependent on the supplier of the needed services and must defer to her to reciprocate for the benefits received lest she lose interest in continuing to provide them. Deference implies not only paying respect to another's superior ability, implicit in asking her help, but also deferring to her wishes in everyday intercourse. Thus, the social interaction among colleagues or in other groups that involves imbalances in social exchange gives rise to differentiation in the power to influence as well as in prestige, which is reflected in a stratified structure of informal status.

The illustration of instrumental assistance in a work group may have left the misleading impression that most social exchange involves instrumental benefits. Much of the social interaction, even among co-workers and still more outside a work situation, is social intercourse engaged in for its own sake. Hechter (1987: 33) states that people often join groups to pursue joint goods or common objectives, and he stresses that their joint achievement and, particularly, the intrinsic gratifications obtained from social associations among fellow members are the sources of group solidarity.⁴

Workers who organize in order to bargain collectively with their employer for higher wages exemplify joint efforts to achieve a common objective. It is in the interest of the group as a whole if workers who devote more energy to and prove more adept in this en-

deavor are allowed to take the leading role in their organizing effort. Thus, superior status based on past services prompts other workers to acknowledge and submit to the leadership of the one who seems to be most effective in making contributions to organizing the nascent union. Informal leadership is legitimated by the social approval of the rest of the group, and this approval is the return for past services and for the future contributions the leader is expected to make to the welfare of the group by helping to organize them.⁵

This fictitious description may well be idealized, but it is not completely inaccurate for the initial stage of workers getting together on their own to organize themselves for joint bargaining. To be sure, it is not applicable to formal positions of leadership, particularly not to the impersonal power their incumbents exercise. Thus, the description is not intended to depict the leadership of large national unions; indeed, it is designed as a contrast to them. Once a union has become a large, formal organization and its leaders have become persons of great power, a handful of workers with a grievance cannot on their own decide upon a course of action if the powerful leader is opposed. All they can do is organize a wildcat strike informally, as workers originally did, but now against both the union leadership and management. The point of this illustration is that the interpersonal power that develops in face-to-face relations is fundamentally different from the impersonal power to dominate large numbers, even in the rare cases when the latter emerged from the former.⁶

Endnotes

1. Cultural theories that explain social patterns in terms of norms and values are prone to commit this tautology. It is the same fallacy as that of psychological explanations of behavior in terms of instincts to engage in such behavior.
2. As indicated by the publication date, this was written long before the women's movement called attention to the implicit bias involved in referring to some unspecified person always by the masculine pronoun instead of using either he/she or even s/he (which I find de-

plorable) or alternating between feminine and masculine pronouns, as I have done in this book.

3. This analysis applies to processes of differentiation in informal status among persons whose formal status is essentially the same.
4. The achievement of joint goods raises the well-known free-rider problem (that persons may benefit from public goods without contributing to their production), which Hechter considers to have solved by distinguishing partly excludable goods from public goods. The former are not available to the entire public but only to group members. His major illustration is that one cannot enjoy the sociability in a group without having become a member and thus a contributor to that sociability. But this solution does not work for instrumental objectives, as indicated by the case next discussed in the text.
5. Workers who fail to contribute to the organizing efforts of the new union would also benefit from its success, which illustrates the criticism I made in the last sentence of the preceding footnote that Hechter's (1987) concept of partial excludability does not solve the freeloader problem for joint instrumental objectives.
6. I am particularly critical of the inference made by conservative social scientists that the elite's domination of society's economy and government is earned as a return for the great contributions they have made to society. It is the counterpart of the assumption that oligopolistic corporations achieved their position in free competition.

References

- Blau, Peter M. 1955. *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press.
- Cook, Karen S., Mary R. Gillmore, and Toshio Tamagishi. 1983. "The Distribution of Power in Exchange Networks." *American Journal of Sociology* 89:275-305.
- Ekeh, Peter P. 1974. *Social Exchange Theory*. Cambridge, Mass.: Harvard University.
- Emerson, Richard M. 1962. "Power-Dependence Relations." *American Sociological Review* 27:31-41.
- Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley: University of California Press.
- Homans, George C. 1961. *Social Behavior*. New York: Harcourt, Brace, & World.

Lévi-Strauss, Claude. 1949. *Les structures élémentaires de la parenté*. Paris: Press Universitaires de France.

Malinowski, Bronislaw. 1961 (1922). *Argonauts of the Western Pacific*. New York: Dutton.

Marsden, Peter V. 1990. "Network Diversity, Sub-structures, and Opportunities for Contact." Pp. 396–410 in C. Calhoun, M. W. Meyer, and W. R. Scott (eds.), *Structures of Power and Constraint*. Cambridge: University Press.

Mauss, Marcel. 1954 (1925). *The Gift*. Glencoe, Ill.: Free Press.

From *Structural Contexts of Opportunities*, pp. 152–163, by Peter Blau. Copyright © 1994 by The University of Chicago Press. Reprinted with permission of The University of Chicago Press. ♦