New Directions in Contemporary Sociological Theory

Edited by
Joseph Berger
and
Morris Zelditch Jr.

4

Development of a Theory of Collective Action: From the Emergence of Norms to AIDS Prevention and the Analysis of Social Structure

Douglas D. Heckathorn

When beginning a research program it is impossible to know where the journey will lead. This chapter describes an ongoing research program of three phases. It began as a study of social norms—the informal standards of conduct that people impose on their family members, friends, and themselves—focusing on special situations in which social norms emerge extremely quickly, even literally over night. The study then broadened to focus on collective action, including not only normative control, but also social movements and broader systems of social cooperation.

The second phase began when, in collaboration with Robert Broadhead, this theory of collective action became the basis for a new form of AIDS-prevention intervention (Broadhead and Heckathorn 1994; Heckathorn et al. 1999). The intervention operates by harnessing peer pressure that creates and strengthens AIDS-prevention norms. Initially tested in several small towns in Connecticut (Broadhead et al. 1998), the intervention has subsequently been implemented in a number of other sites, including Yaroslavl, Russia (Sergeyev et al. 1999); and the principles upon which the intervention is based have been adapted to address other public health issues, including controlling high blood pressure and increasing adherence to AIDS therapy (Broadhead et al. forthcoming).

The third phase developed as an unanticipated by-product of the AIDS-prevention intervention, leading both to improved means for sampling hard-to-reach populations, such as drug injectors and the homeless, and for studying social structure, which, following Pareto, is conceived as structured patterns of affiliation.

COLLECTIVE ACTION

The research program began as a study of the emergence of norms (Heckathorn 1988). A problem faced in any such study is that the origins of

most norms lie deep in the inaccessible past. For example, all societies have norms regulating theft, interpersonal violence, and deception. Though these norms are subject to contextual variations, as when higher-status people are granted privileges denied their social inferiors, and they change over time, as when new technologies such as the Internet create new opportunities for theft and deception, changes tend to be slow and are heavily influenced by established conventions. Examples of pristine norm emergence are rare.

As I was searching for contexts where norm emergence could be studied, I recalled a story told by a friend's father (Will Brothers) concerning his experiences as a drill instructor in a U.S. army boot camp during World War II. Discipline in U.S. boot camps is based on collective punishment; if one recruit breaks a rule, all recruits in the barracks are punished. Recruits whose violations have provoked the punishment of their peers are sometimes beaten in an institution known as the "blanket party." The speed with which these norms emerge is remarkable—they literally emerge and are vigorously enforced overnight. Whereas recruits might otherwise enjoy watching peers challenge the authority of the drill instructor, much as class clowns earn the esteem of their peers by tormenting teachers, norms prohibiting challenges to authority emerge quickly and are fiercely enforced in boot camps. Further investigation revealed that use of collective punishment in boot camps has been discussed both by sociologists (Gilham 1982) and by playwrights (e.g., Neil Simon in *Biloxi Blues*).

Collective Sanctions and Norm Emergence

Norm emergence based on collective punishment is explicable as purposive behavior. When the actions of others can provoke collective punishment, this creates a *regulatory interest*, that is, an incentive to regulate the actions. Described in economic terms, the action generates *externalities* (i.e., costs resulting from the collective punishment triggered by the action). Thus, the creation and enforcement of norms occurs when individuals act based on regulatory interests, which in turn are based on externalities. Similarly, norms can be created by collective rewards, as when members of a team earn rewards based on the team's success. The recognition that collective sanctions, either collective punishment or collective rewards, could trigger norm creation suggested that settings in which these were used could serve as a real-world laboratory for studying norm emergence.

Note that here norms are not seen merely as constraints on behavior, but rather as something that social actors *do*, that is, as a form of social action. Some actions affect only the actor. Other actions have consequences for others, especially the exercise of social influence, in which the aim of the action is to affect others' behavior. Social influence can take many forms, including persuasion intended to make others want to act in the intended way, incentives such as promises and threats, and even physical compulsion. By whatever means, individuals have considerable ability to influence one another's behavior. This is one of the

consequences of mutual interdependence, that all of us are dependent in myriad ways upon those around us. An especially important form of social influence is based on social approval, a form of sanction that is important in all human societies. Though exercise of some social influence is idiosyncratic, much social influence is based on widely shared regulatory interests. When a widely shared regulatory interest leads to a consistent and coordinated exercise of social influence—that is, to *collective action*—a social norm can be said to exist. Thus, social norms are a form of collective social action. For example, the preference for physical safety is the regulatory interest underlying norms limiting violence, and the preference for security in one's possessions is the regulatory interest underlying norms prohibiting theft. These are regulatory interests that are, in a sense, built into the human condition, so all societies have norms limiting violence and defining property rights. The origins of these norms are therefore lost in antiquity. In contrast, when regulatory interests are created by collective sanctions, either collective reward or collective punishment, this creates a sort of real-world laboratory for the study of norm emergence.

The next step in the project was a survey of the literature on collective sanctions, which turned out to be voluminous. However, it was also fragmented, divided into a number of unrelated literatures. For example, there is a large anthropological and historical literature. Most traditional legal systems are based on a principle of corporate responsibility, in which the extended family is jointly responsible for each member's actions. In nineteenth-century Albania, if a person from one village killed someone from another's village, anyone from the victim's village was entitled to kill anyone from the culprit's village within twenty-four hours of the offence, and to kill anyone from the extended family of the culprit within forty-eight hours (Heckathorn 1988). Obviously, this strengthened incentives to regulate behavior within one's village and within one's household. Similarly, in traditional China, capital punishment for especially serious offences was extended to the culprit's father, brothers, and son; and women from the family were sold into slavery. As in the U.S. boot camp, the effect of such a system of collective sanctions was to strengthen the incentives of family members to regulate one another. Other literatures focus on collective sanctions in schools, firms, and prisons. Each documents cases in which collective sanctions create compliance norms, thereby co-opting informal norms, making them into an extension of the system of legal or organizational authority. When collective sanctions are successful, the norms that reinforce the authority of the agent controlling the sanction are termed *compliance norms*.

However, collective sanctions are not always successful, as illustrated by the Algerian revolution. According to Heggoy (1972:235), the French army overreacted to terrorism in Algiers:

> the terrorist cells were dismantled and most of the members were arrested or killed. The strategic victory, however, belonged to the nationalists, who reaped immense political gains from the high-handed military tactics. [The French] created isolated

Algerian ghettos whose occupants grew increasingly united in their hatred of France. . . . The difference between Algerians and Europeans living in Algeria became markedly clearer. . . . This development forced the two communities to drift further and further apart. The nationalists capitalized on the social fault thus created and undertook the leadership of the Algerian population as a whole.

In essence, the nationalists used terror tactics to provoke the French colonial government to punish collectively and withhold collective rewards from the Algerian middle class and native Algerians in general. The result was to increase the incentives for previously procolonial Algerians to throw in their lot with the nationalists. This use of polarizing tactics based on terror is a time-honored tactic for fomenting revolution. Classic examples are the Sicarri and Zealots who opposed the Roman Empire's rule of Judea starting in the year A.D. 4. Public assassinations of Roman soldiers provoked the Romans into reprisals against civilians, which in turn increased public opposition to the Romans, eventually resulting in revolt (see Rapoport 1984).

Thus, collective sanctions create *ambivalent* incentives: incentives both to create and enforce compliance norms, and also opposite incentives to attack the source of the sanctions, to destroy its ability to dispense collective punishment or withhold collective rewards.

A formal mathematical model was constructed to explain when collective sanctions would either result in compliance norms or provoke revolt and to unify the diverse literatures in which collective sanctions were analyzed (see Heckathorn 1988). Consistent with the conception of norm emergence and enforcement as a form of purposive action, the first step in constructing the model was to define the actors composing the system. This included both the group of individuals subject to the collective sanction and the agent controlling the sanction. The second step was to define the options available to each actor and the costs of choosing each option. The agent was assumed to have issued some form of dictate, it had a specified ability to monitor behavior within the group to determine the degree of compliance with the dictate, and it could choose a threshold level of noncompliance that would trigger sanctions of specified severity. No prejudgment was made regarding either the extent of the agent's monitoring capacity or the severity of the collective sanction. These served as variables in the model, because one of the aims was to study the effects of changes in these terms. The members of the group made two interrelated choices. First, they were assumed to be able to choose whether to comply with the dictate. Second, they could choose to employ whatever social influence they possessed to encourage others to comply, thereby supporting a compliance norm, they could revolt, by seeking to deny the agent the ability to dispense collective punishment or withhold collective rewards, or they could do neither and remain passive. No prejudgment was made regarding the costs of complying with the agent's dictate, the extent to which each individual could control the behavior of others (i.e., group cohesion), the costs of exercising that control, the vulnerability of the agent to revolt, the cost of participating in a revolt, and the size of the group subject to col-

lective sanctions. These also served as variables in the system, because the aim was to understand the consequences of changes in such factors as group size and cohesion. Third, mathematical analyses were carried out to determine how each of the model's variables affected the strength of the incentive to either create compliance norms or revolt.

The analyses revealed that controlling a group through collective incentives is rather like walking a tightrope. For the same factors that encourage members to choose compliance over passivity, such as employing a strong collective sanction, may also induce them to revolt. Similarly, moderately stringent demands for group compliance are most effective in producing compliance norms, because too stringent demands for group compliance tend to provoke revolt because sanctions come to be seen as inevitable, and too lenient demands cause the agent's dictate to be ignored. Another conclusion was that moderately cohesive groups are most effectively controlled by collective sanctions, because highly cohesive groups tend to revolt, and atomized groups lack the capacity to create effective compliance norms. In general, collective sanctions are most effective when used by a strong agent to control a small and relatively cohesive group, as in the example of the boot camp. Collective sanctions fail when used by vulnerable agents against a large and dispersed group, as in the example of the French in Algeria.

The analyses also lead to the recognition of institutional means that have evolved to reduce the revolt-inducing potential of collective sanctions. For example, when criminal punishment extends from the culprit to the culprit's family, the effect is not merely to strengthen incentives to encourage family members to behave lawfully, but also to encourage family members to assist and even join in criminal pursuits. To invert the usual phrase: if one will do the time, one might as well do the crime. To avoid such problems, the Chinese legal system permitted parents to divorce a child by making a payment to the emperor in anticipation of any damage the child might do. This helped to ensure that an incorrigible family member would not pull the entire family into complicity with or active participation in wrongdoing. The question of the effectiveness of the Israeli policy of blowing up the family homes of accused terrorists has been much debated. Whereas it may deter some measure of terrorism, it may also foster hostility toward the Israeli government, and thus may increase rather than reduce resistance. According to this model, the ultimate effectiveness of this policy would depend upon the stringency of sanctions, the perceived vulnerability of the Israeli government, and the cohesiveness of the Arab families involved, especially whether the person whose actions triggered sanctions continued to live in the household, and hence was under the potential control of household members.

Network-Mediated Social Control

The literature on collective sanctions focuses on special settings, including total institutions and revolutionary systems. However, after studying these sanctions, it became apparent that they are less rare and exotic than had been previously recognized (Heckathorn 1990). For virtually all individuals are members

of groups with which they are interdependent. These include groups of family members, friends, neighbors, coworkers, and others with whom they interact regularly. To the extent that members of a group are interdependent, events that impact on any individual have consequences that extend to other group members. For example, when one person is promoted on the job or fired, the sanction spills over and affects family members and friends. Therefore, except in the limited case of social isolates, almost all social sanctions targeted at an individual generate collective rewards or punishments that impinge on his or her primary group. Imprisonment is an example of a punishment that spills over to others. It is not merely a personal calamity; it frequently drives whole families into poverty. Similarly, rewards spill over when a family's major breadwinner earns an important promotion because it improves the entire family's circumstances. Due to the spillover of rewards and punishments from individuals to others, social sanctions are virtually never individualized. Instead, they give rise to collective rewards or collective punishments.

Given that most social sanctioning includes both an individual and a collective component, behavioral compliance can arise from either of two theoretically distinguishable sources (see figure 4.1). First, it can arise from individual sanction-based control directed at an actor by an agent such as a teacher, parent, neighbor, or AIDS-prevention counselor. For example, an agent may target an actor with the promise of a reward or a threat of punishment. The result is a dyadic relation of the sort presumed in most analyses of influence relations. This is represented by the hollow arrow in figure 4.1. Second, compliance can also arise from network-mediated control, as when students obey teachers because punishment administered by the school would be augmented by parents, or when workers hold onto
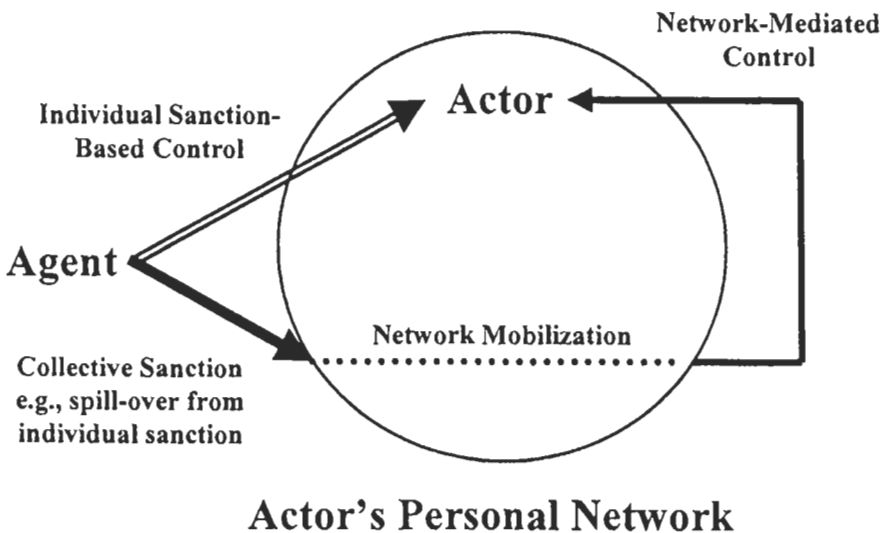

Actor's Personal Network
Figure 4.1. Network Embeddedness of Social Influence

disagreeable jobs because unemployment would inflict hardship on their families. In these cases, control occurs through a two-step process. First, the actor's group is promised a collective reward or threatened with a collective punishment based on whether the actor complies. Second, the group responds to that incentive by controlling the actor. In this way, the agent's influence is amplified through the group in which the target of control is embedded. This two-step process is represented by the solid arrows in figure 4.1.

Control based on individual sanctions works by altering peoples' *inclinations*, that is, their preferences regarding their own personal behavior (see table 4.1). It does this by using what may be termed *primary incentives*, such as performance-specific rewards or punishments. In contrast, network-mediated social control works by altering peoples' *regulatory interests*, that is, their preferences regarding how others behave. Network-mediated social control does this by using what may be termed *secondary incentives*, such as rewards or punishments based on the performance of peers. In other words, secondary sanctions create externalities—costs or benefits that are contingent on the behavior of others. Thus, secondary sanctions alter the externalities upon which regulatory interests are based.

Building on the previous analysis of collective sanctions, a formal model was constructed to analyze systems that include both primary and secondary incentives. In many respects, the setup for the model resembled the previous model. The system included both an agent and a group subject to the agent's sanctions. However, the sanction had not merely a collective sanction, but also individual

---

Table 4.1. Network-Mediated Control: Basic Terms

Two Type of Preferences

Inclination—A person's preferences regarding his or her own personal behavior.
Regulatory Interest—A person's preferences regarding how others should behave.

Two Types of Incentives

Primary Incentive—A reward or punishment that targets a specific individual based on his or her performance. If sufficiently strong, this changes the individual's inclinations.
Secondary Incentive—A reward or punishment administered to a group based on the performance of one of the group's members. If sufficiently strong, this changes the members' regulatory interests.

Two Types of Behavioral Control

Individual Sanction-Based Control—Behavioral control based on primary incentives that alter inclinations, thereby motivating compliance.

This is the type of control typically assumed in studies of organizational power and in theories of deviance and crime.

Network-Mediated Control—Behavioral control in which a secondary incentive alters group members' regulatory interests and thereby motivates them to exert control over one another.

According to the theory of network-mediated control, this is the type of social control that accounts for most compliance within organizations and communities.

sanctions, sanctions that could be targeted to individuals. Consistent with the recognition that when members of a group are interdependent, something that affects one individual thereby affects all of them, each individual sanction was assumed to spill over to other group members, thereby generating a collective sanction. The magnitude of that spillover served as a variable in the model. Members of the group could choose among multiple options. First, as in the previous model, they could choose whether to comply with the agent's dictates. Second, again as in the previous model, they could choose whether to use their influence to pressure others into compliance, or they could remain passive. Because the focus was on the relationship between primary and secondary incentives, the option of revolt was dropped from the model, and replaced with the option to oppose the emergence of compliance norms. This option involved exercising social influence to prevent others from creating and enforcing compliance norms. Variables in this model included the strength of the individualized sanction, the amount of spillover, which in turn determines the strength of the collective sanction, the cost of exercising compliant control (i.e., the cost of participating in enforcing compliance norms), the strength of the control, the cost of oppositional control, and the strength of that form of control. The interpretation of this basic model was broader than in the previous collective sanction analysis. For the agent could represent not merely purposive agents, such as supervisors, teachers, or police officers, but also a natural process, as in what game theorists term a "game against nature." For natural processes can also have both individualized and collective effects. For example, fire can spread to threaten neighbors' homes or fields, and unsanitary conditions can trigger epidemics. Thus, the "agent" could represent a collective bad, such as fire or disease, or a collective good, such as agricultural land or a fishery. With this expanded interpretation, the analysis was broadened to apply to collective action systems in general. The analysis of this model was based on a view of actors as purposive in the sense that their behavior was incentive driven (Heckathorn 1990).

A central conclusion from the theory was that secondary incentives could be both more efficient and more effective than primary incentives (Heckathorn 1990). This occurs when the means for intragroup control are both cheap and effective, as occurs in cohesive groups where peer approval is an important sanction. These conditions fit many primary groups such as family and friendship groups. In the case of individualized sanctions, compliance costs are internal. The targeted actor either complies or refuses to comply with the directive, and thereby bears whatever costs are involved. In contrast, in the case of secondary incentives, compliance costs are external. The targeted actor either succeeds or fails in inducing another actor to comply, so the costs of compliance are borne by that other actor, and the targeted actor bears merely the typically minimal cost of exercising social influence. It is usually easier to tell others what to do than to do it oneself. In extreme cases, one might urge others to endure great suffering to avoid minor inconvenience or embarrassment to one-

self. As a result of the externalization of the costs of compliance, secondary incentives can produce a state termed "overcontrol" (Heckathorn 1990, 1991), in which compliance occurs even though its costs exceed those of noncompliance. This produces a social dilemma in which collective gains result not from strengthening but from weakening normative controls (Heckathorn 1990). To the extent that this occurs, secondary incentives can be so effective that they can produce superoptimal levels of compliance, and thereby create "zealots" in Coleman's (1990) sense.

A second reason for the exceptional effectiveness of secondary incentives concerns monitoring. Monitoring for primary incentives is undertaken by the agent controlling the sanctions. Police, teachers, and supervisors typically can observe only a small portion of behavior, so monitoring is difficult when activities cannot be geographically confined. In contrast, secondary incentives operate through peer influence, and peers tend to be far more effective monitors of behavior (Heckathorn 1990).

A third reason for the exceptional effectiveness of secondary incentives concerns what is termed the "hidden cost of reward." Material rewards may undermine intrinsic motivation when they are framed as "pay" rather than as "recognition" for achievement (Deci and Ryan 1985). This creates a dilemma for organizations that rely on primary incentives, because if their ability to reward symbolically is limited, they must rely on material rewards. In contrast, secondary incentives harness peer pressure, so they rely on nonmaterial rewards such as peer approval to secure compliance. Thus, whatever intrinsic motivation exists will be preserved, and even strengthened, because of peer support. Secondary incentives do, however, present a potential problem. If the secondary incentives employ material rewards, might they undermine intrinsic motivations to engage in peer influence? That is, might they weaken preexisting peer norms? According to the formal model, this does not occur because the exercise of peer influence entails a *commitment*. When a person urges peers to act in a certain way, that person publicly affirms the special value of that behavior. If that person then attempts to retreat from the commitment, he or she risks appearing opportunistic or hypocritical. Thus, inducing individuals to affirm publicly the value of acts serves as a means to strengthen their commitment to them. Indeed, public affirmations of commitment to particular ways of behaving are a fundamental and powerful social mechanism for creating and maintaining social cohesion.

The implication of this analysis is that treating relationships of power and authority as a set of dyadic relationships between superordinate and subordinate is a mistake, for reactions to sanctions depend on the networks in which the individuals are embedded. These groups can either amplify the effects of sanctions—and thereby enhance authority—or counter the effects of the sanctions—and thereby undermine the authority. Furthermore, the analyses suggest that much legal and organizational control derives not from the dyadic relationship between controller and target of control, but rather from the circuitous

process described as network-mediated control. The implication is that some of the literature on legal and organizational control must be rethought to take into account the embedding of control relationships within larger social systems.

This theory of collective action resolved a problem that afflicted Olson's (1965) theory. According to Olson, the manner in which collective action is organized depends on group size. In small groups, each individual's stake in collective action is so strong that participants are not tempted to free ride. Each individual's net gain from his or her own participation exceeds its cost. These groups are termed "privileged." In medium sized groups, a temptation to free ride exists that is controlled through strategic interaction, in which each individual's contribution is conditional upon the contribution of others. Thus, collective action is based on reciprocity. In large groups Olson argued that collective action required selective incentives, such as punishment for noncontributors (e.g., criminal penalties for tax evasion) or rewards for contributors. Individuals then contribute, not because they value the collective good, but merely because of the sanctions. Thus, contribution is a *by-product of the sanctions*. Thus, the decision to contribute to the collective good is decoupled from any valuation of the collective good. This *by-product theory of collective action* has been criticized, because by explaining too much, it explains too little. It explains contribution in terms of selective incentives, but it does not explain the origins of those incentives.

A system of selective incentives, including a normative system, is itself a collective good. For the norms confer benefits even on those who did not contribute to their production, that is, those who did not participate in norm enforcement. Therefore, a free-rider problem arises, termed the *second-order free-rider problem* (for an analysis, see Heckathorn 1989). The problem with Olson's by-product theory is that it does not explain how this problem is resolved. Yet, the above model provides a resolution. Recall that in the model, regulatory interests govern the exercise of social influence, and the cost of exercising that influence is generally low. It is almost always easier to tell someone else what to do, than to do it oneself. When a collective good is valued, this creates a regulatory interest, an incentive to encourage others to contribute to its production. Regulatory interests are strengthened based on the number of persons one can control, the effectiveness of that control, and valuation of the outcome of that control. Though this regulatory interest can be weak in large groups, if the costs of exercising influence are correspondingly low, even weak regulatory interests can shape the exercise of influence. Therefore, in large groups, where the significance of any individual's contribution is small, norms mandating contribution can nonetheless arise. Thus, regulatory interests provide the link that was missing in Olson's theory, between valuation of the collective good, and selective incentives.

The model was further extended to analyze the effects of group heterogeneity on norms (Heckathorn 1993). In a series of papers, Marwell, Oliver, and associates (Oliver, Marwell, and Teixeira 1985) had argued that contrary to conventional wisdom (Olson 1965), increases in both group size and heterogeneity

promote collective action. This occurred, they argued, because the larger and more diverse a group, the greater would be the number of individuals with an especially strong interest in promoting collective action. This group would then serve as a "critical mass" that would trigger the emergence of collective action. This analysis considered only a single way in which collective action could be organized, voluntary cooperation in which each individual chooses independently whether to contribute to the collective endeavor. They therefore ignored selective incentives.

To evaluate the Marwell-Oliver analysis, the network-mediated control model was extended to include three forms of heterogeneity, variation in the extent to which the collective good was valued, variations in the cost of contributing to the collective good, and variations in the size of the contribution the individual could make. More specifically, each of these terms was assumed to be normally distributed. For example, whereas in the previous analysis (Heckathorn 1990), all group members were assumed to value the collective good equally, in this new analysis (Heckathorn 1993), valuations varied around a specified mean and standard deviation.

The conclusions from this analysis differed from those of Marwell and Oliver. For the analysis (Heckathorn 1993) showed that depending on the circumstances, heterogeneity can either promote collective action, or it can cause the group to fragment into mutually antagonistic factions. The latter can occur when what is for some a collective good is for others a collective bad, or when costs of contribution vary. Such cases are common in real-world collective action problems. For example, when environmentalists promote regulations to protect what they see as valuable and fragile ecosystems, the affected industries often complain about loss of jobs. The analysis further showed that polarization is especially likely when collective action is organized through selective incentives. For selective incentives compel even those who lack any interest in the collective good to contribute, and thereby provide those individuals with an incentive to mobilize in opposition. This was an issue Marwell and Oliver did not consider, because they considered only voluntary contributions. However, public policies reflect recognition of the potentially divisive nature of selective incentives. Politicians are frequently reluctant to support use of public funds for controversial programs. For example, in New York, state-sanctioned needle exchanges do not receive public funds. They operate through private donations. This ensures that individuals who oppose these exchanges will not be taxed to support a program they do not support, and thereby weakens their incentive to mobilize in opposition to the exchanges.

INSTITUTIONAL DESIGN

The second phase of this research program was based on a chance meeting. After having concluded that collective incentives were potentially more

effective than the individualized incentives upon which most of the literature on organizational and social control was based, I began seeking ways to test this proposition. At that time, I met a medical sociologist, Robert Broadhead, and we began discussing a process evaluation he had conducted of AIDS-prevention projects targeting active injection drug users in San Francisco and New York. He was struck by an apparent paradox. The interventions worked quite well in reducing AIDS risk behavior such as sharing syringes. Yet the outreach workers who carried out the community education performed poorly. Broadhead suggested that the solution to this paradox lies in the role of injectors—they played a highly active role in passing along what they had learned from the outreach workers to their peers, thereby amplifying the impact of the intervention.

Over a series of lunches, Broadhead and I discussed the question of whether my work on collective action could be used as the basis for a new form of AIDS-prevention intervention that would create and strengthen prevention norms among injectors. More specifically, the idea was to base the intervention on secondary rather than primary incentives, so that the target of the intervention, active injectors, would take over many of the roles typically performed by full-time professionals. Broadhead's evaluation of outreach worker–based interventions demonstrated that injectors were capable of playing an energetic and constructive role in a prevention intervention, so the idea seemed promising.

Some background on traditional approaches to AID prevention may be useful. Traditional AIDS-prevention efforts for injection drug users (IDUs) have been based on a "provider-client" model called "street-based outreach" (Brown and Beschner 1993). The model involves hiring a small number of community members, usually ex-addicts or people with street credentials, to contact and work with members of their own community as clients. They do this by going into neighborhoods as "outreach workers" (OWs) to distribute AIDS-prevention materials and information, and to recruit injection drug users (IDUs) to various programs and services, including research interviews conducted by social scientists.

Outreach projects operate under conditions that cause hierarchical supervision to break down. First, moral hazard problems abound because of the opportunities to gain illicitly from working in active drug scenes. For example, Broadhead and Fox (1990) reported cases where outreach workers used their jobs as a cover for drug dealing and fencing stolen goods. Second, adverse selection problems are severe. Being streetwise is an essential qualification for an OW. Such individuals are generally accomplished hustlers, so distinguishing those with a sincere desire to work to prevent AIDS from those who simply wish to hustle a project is virtually impossible; projects can only find out they have been conned *ex post* (Broadhead and Heckathorn 1994). Finally, monitoring of performance is necessarily limited because AIDS-prevention outreach occurs on the street, in single-room-

occupancy hotels, public housing projects, public parks and so on. In sum, moral hazard and adverse selection problems are unusually severe, and monitoring of OWs' performance is limited. The result is an array of organizational problems that invite and allow poor performance by OWs to go on virtually unnoticed, and that push outreach projects toward inertia (Broadhead and Heckathorn 1994).

In contrast, IDUs responded impressively to the outreach services they received; IDUs adopted many risk reduction measures, and they volunteered and substantially augmented the efforts of OWs. Put simply, IDUs went well beyond the role of being mere clients; their response to traditional outreach was far more robust and far-reaching than were the efforts of outreach projects themselves.

In light of users' responsiveness, an AIDS-prevention intervention that relied on IDUs as active collaborators seemed feasible. Such an intervention would contrast with the traditional model that turns IDUs into clients of, and makes them dependent on, paid staffs of OWs. The model would also draw upon and strengthen the sharing rituals and norms of reciprocity that already underlie and sustain drug user networks in the first place (Preble and Casey 1969). It would do so by enhancing the mutual opportunities and incentives for IDUs to work with their own peers, and to invest themselves in their own intervention.

A Peer-Driven Intervention

Based on the theory of network-mediated social control (Heckathorn 1990), we designed such a model, called a "Peer-Driven Intervention" (PDI), that began operating in March, 1994 in eastern Connecticut (Broadhead and Heckathorn 1994). The PDI was compared to a traditional outreach intervention operating in a separate but demographically similar community. The PDI uses secondary incentives to harness the potentially enormous power of peer-pressure as a means for altering behavior.

The PDI design employed a two-step process. First, the essential activities of traditional outreach were identified. Second, a structure of secondary incentives was implemented that offers IDUs recognition and modest material rewards for encouraging their drug-using peers to carry out prevention activities in their own community.

The first task of traditional outreach is recruiting IDUs into prevention programs. As in traditional programs, the nexus of the PDI is a storefront facility within which AIDS testing and counseling services are offered, as well as risk reduction education and materials such as bleach kits and condoms. In the PDI, IDUs are motivated to recruit other users for the above services via a coupon system: for each IDU recruited bearing a coupon, the user who recruited him or her receives a monetary reward. Only modest rewards are required, because the cost involved in exercising influence over peers is small,

Table 4.2.   **Incentive Structures in Two Types of Interventions**

Task	Traditional Outreach Intervention	Peer-Driven Intervention
Recruit users for interviews, AIDS education, and HIV antibody test counseling.	Outreach Workers (OWs) are assigned the task of accessing and recruiting injection drug users (IDUs) in the community.	Each IDU is given three recruitment coupons. She is then told that if a peer she has recruited comes to the program for an interview with her coupon, she will be rewarded.
	Continued employment is conditional on satisfactory job performance. The reward structure relies on primary incentives to motivate OWs to provide services to their IDU clients.	The monetary reward structure is mixed, because it combines a primary incentive (the reward for being interviewed), and secondary incentives (the rewards for recruiting peers). Project staff also strongly emphasize altruistic motivations for recruiting peers.
Educate users about AIDS, harm reduction, and other health measures.	OWs are assigned the task of educating users in the community.	Each IDU-recruiter who educates a peer about AIDS prevention, as measured by a brief eight-item knowledge test administered before each interview, earns an additional reward. Because a reward is earned from eliciting a positive response from a peer, education is motivated by a secondary incentive. Project staff also strongly emphasize altruistic motivations for educating peers.
	Continued employment is conditional on satisfactory job performance. The reward structure relies on primary incentives to motivate OWs to provide education to their IDU clients.	
Distribute AIDS prevention materials.	OWs are assigned the task of distributing prevention materials in the community.	Before leaving the storefront, each IDU is given prevention materials for her personal use and for distribution to peers. Prevention materials are valued by users, so no exogenous incentives are required for their distribution.

and there now exists widespread concern about AIDS within the injection community—which is to say that *regulatory interests among IDUs concerning AIDS prevention already exist.* Furthermore, recruiting and educating peers involves a public commitment to AIDS prevention that reinforces those regulatory interests. Each recruit, in turn, is also given a small number of coupons to recruit still other IDUs within their drug-using network. Thus, the mechanism co-opts user networks to serve as a medium to recruit further IDUs.

This approach has several advantages. First, it puts the burden of identifying recruits on those with the best current information: active users. Second, the PDI's pay-for-performance design recognizes and rewards the most productive recruiters. As a result, subjects are rewarded in direct proportion to the success of their recruitment efforts, and those who recruit no one receive nothing.

Third, a PDI offers a built-in accommodation to the cultural diversity in the user population: with IDUs accessing their peers, the recruitment effort is always couched in culturally appropriate terms for each user subgroup. Thus, built into a PDI is a performance-based reward system that continuously adapts to cultural and other subgroup differences.

Another central task of outreach is distributing AIDS-prevention information. Traditional programs educate IDUs both in the field, and at a storefront, van, or similar space. In a PDI, IDUs are given the training, and incentives to educate their peers in the community. The extent to which IDUs pass on information to those they recruit is measured through questions added to standard interview schedules, and the reward to the recruiter depends on the knowledge of the recruit. This approach has several advantages. First, it puts the responsibility for educating IDUs on those who are most likely to be influential: their peers. Second, it entails considerable repetition. Subjects are first educated by their peer-recruiter, then by project staff, then subjects rehearse what they have learned when educating and recruiting several of their peers. Third, its pay-for-performance design recognizes and rewards the most effective educators.

When the PDI was compared with the traditional outreach project, several findings were notable. The data were based on an initial interview and one six-month follow-up interview that assesses health status and behaviors that put people at risk for AIDS, sexually transmitted diseases, and other drug and sex-related health problems. The first finding was that IDUs can indeed be motivated by secondary incentives to recruit their peers. When compared to the traditional outreach control site, the PDI succeeded in recruiting 50 percent more IDUs during the first calendar year of operation, as well as a sample that was more heterogeneous in drug preference, and drawn from a wider geographic area (Broadhead et al. 1995).

Second, IDUs can be motivated, and are able, to provide effective AIDS-prevention education to their peers in the community. They respond favorably, and sometimes enthusiastically, to the opportunity to acquire potentially life-saving information and share that information with their peers. Some recruiters reported that they went over the lesson several times with their recruits. In anticipation of their role as a peer educator, some IDUs have even taken notes during the education they received from project staff, and some have also called the project to ask for further clarification. Scores on the knowledge assessment test show that peers can educate as effectively as professional OWs, and after IDUs have recruited and educated several peers, recruiters' own knowledge level significantly exceeds that of subjects in the control intervention (Broadhead et al. 1995).

Third, and most significantly, the PDI produces reductions in AIDS risk behaviors, including reductions in estimated injection frequency and syringe sharing. Figure 4.2 summarizes the effects of PDIs on syringe sharing behaviors in New London (Heckathorn et al. 1999), Middletown and Meriden, Connecticut (Broadhead et al. 1998), and Yaroslavl, Russia (Sergeyev et al. 1999). These are compared with the one- and two-year impact results from the project's control
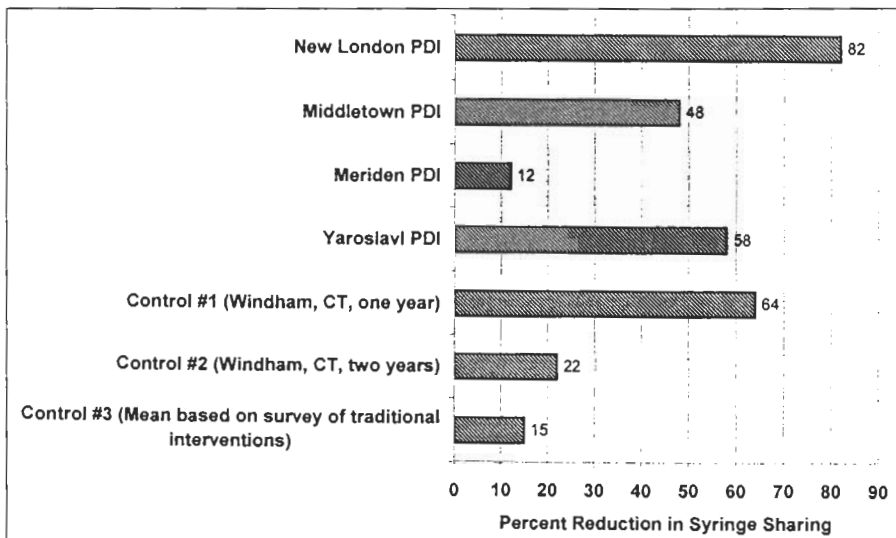
Figure 4.2.   Changes in Syringe Sharing by Intervention

site, a traditional intervention implemented in Windham, Connecticut. PDI impact results can also be compared with impact results from the dozens of traditional interventions that have been implemented and evaluated in the U.S. Needle and Coyle (1997) surveyed studies assessing these traditional outreach interventions, and they provide a national benchmark against which PDI impact can be compared.

All interventions reduced syringe sharing, but to varying degrees. Three of the four PDIs produced substantial reductions, from 48 percent in Middletown, 58 percent in Yaroslavl, to 82 percent in New London. The fourth PDI's impact was substantially lower, only 12 percent. Though further research will be required to confirm this hypothesis, this difference appeared to result from a "proximity effect" (Broadhead et al. 1998:52). The three more effective PDIs were located near the center of each city's drug scene. In contrast, subjects in Meriden had less ready access to the PDI, in most cases requiring a twenty-minute bus ride, and then walking a mile. This disadvantaged them in several ways, for example, making it difficult for them to return to the intervention's storefront for prevention materials during which prevention messages could be reinforced.

The impact results from the three controls are also variable. In a survey of dozens of standard interventions conducted by Needle and Coyle (1997), reductions in syringe sharing ranged from 13 percent to 43 percent, with a mean of 15 percent. The one-year and two-year impact results from our project's control site were also variable, with substantially stronger reductions during the first year than for two years. The reasons for this reduction remains unclear, though it may be related to a volunteerism effect, in which the most educable

and tractable subjects made themselves available for recruitment by the traditional intervention's professional outreach workers, and after the first year, this especially tractable pool of subjects had been exhausted.

Overall, when the PDI impact results are compared with that from traditional interventions, most of the PDIs perform respectably. Specifically, three of the four PDIs not only exceeded the mean for previously published evaluations of standard interventions, but also exceeded the maximum reduction reported in these studies. These benchmarks were also exceeded during the control intervention's first year, though these results were not sustained. The variability in these results suggests the need for more research to better identify the determinants of intervention impact, but what is clear is that under suitable conditions the PDI can serve as an effective means for reducing AIDS risk behavior.

Fourth and finally, because of its greatly reduced reliance on professional staff, the PDI is far less expensive than traditional outreach. At our control site, the full-time salaries and fringe benefits paid to the OWs resulted in recruitment and education costs that averaged $623 per recruit, as compared to only $14 per recruit in the PDI. Hence there is more than a forty-fold differential in cost. This is an important issue for any public health intervention, because budgets for public health are always limited.

Our approach to AIDS prevention was supported in a report from the Institute of Medicine (1995) of the National Academy of Sciences. The report's aim was "assessing the social and behavioral science base for HIV/AIDS prevention and intervention," and it was "targeted primarily to policy-makers who will be making decisions for the HIV/AIDS research agenda in the next decade." The PDI was described as "the state of the art of preventive intervention."

SAMPLING AND SOCIAL STRUCTURE

The third phase of this research program began as an unanticipated spin-off of the AIDS project. Recall that the PDI's recruitment mechanism was based on a chain-referral process, where the initial respondents (i.e., the "seeds") each recruited several peers, who each recruited several more peers, and so forth as the interventions expanded wave by wave. When designing the intervention, Broadhead and I had concluded that recruitment quotas were necessary to prevent the emergence of professional recruiters, who might then seek to monopolize recruitment rights and even battle for turf. The question was, how should the quota be established? If the quota was too small, recruitment would die out; and if too large, professional recruiters might emerge.

To provide a principled means for exploring these questions, I constructed a mathematical model with which to simulate the recruitment process (Heckathorn 1997). For simplicity in constructing the model, the population was assumed to be indefinitely large, to correspond to an intervention drawing

from a large population of potential participants. This model provided the means for simulating the effects of alternative quotas, including exploring the ability of the recruitment process to reach groups who were socially distant from the seeds with which recruitment began. That is, the simulated population was divided into distinct groups, corresponding to race/ethnicity or other social categories, with differential connections among groups. Based on these analyses, the quota was ultimately set at three recruits per respondent.

These simulations produced an unexpected result, the finding that after a modest number of recruitment waves, the composition of the sample became the same, irrespective of the choice of initial seeds. That is, whether all the seeds were drawn from a single category, or dispersed among all the categories, the sample composition ultimately converged upon a single equilibrium. Furthermore, unless population groups were nearly totally socially isolated from one another, this finding was unaffected by the network structure of the population. A bit of checking confirmed that when the recruitment process was modeled as a form of stochastic process called a "Markov chain" (Kemeny and Snell 1960), this result corresponded to a well-known theorem, the law of large numbers for regular Markov chains.

The equilibrium result suggested a new extension of the research program, because it suggested that the AIDS intervention's recruitment process could serve as a sampling process that was *reliable*. That is, were the sampling to be repeated (e.g., starting with a different set of seeds), the sample composition would be the same. Further analysis addressed the issue of *validity*, that is, whether the sample would correspond to the population from which the sample was drawn.

This phase of the project began as an effort to develop better means for sampling hard-to-reach populations such as injectors, and evolved into a new approach to studying social structure. Let us first consider sampling. Great attention has been devoted recently to the problems involved in sampling hidden populations because of two recent events, the AIDS epidemic (Watters and Biernacki 1989; Laumann et al. 1989) and decreases in the accuracy of the U.S. census (Brown et al. 1999). Efforts to address both problems have focused attention on problems in sampling hidden populations. The primary focus has been on injection drug users, men who have sex with men, and the homeless.

Given its small size, using traditional methods to sample a hidden population would be prohibitively expensive. Furthermore, when a hidden population has privacy concerns, it cannot be reached by methods such as household surveys or random digit dialing, nor can these methods reach those with unstable living arrangements as when several families live in an apartment although only one's name appears on the lease (Sudman and Kalton 1986).

Three methods currently dominate studies of hidden populations. First, location sampling involves identifying locations where members of the population can be found, and then deploying interviewers. A problem is that location sampling is practical only for locations that are large and public. However, such

large public scenes tend not to draw a representative sample of any hidden population. For example, not all injectors buy their drugs on the street.

Institutional samples are a second method for sampling hidden populations. In the case of drug injectors, samples are drawn from drug-treatment programs and prisons. Here, the researcher relies on the institution to draw the sample. The problem is that only a select group of subjects enter drug treatment programs, prisons, and other institutional settings.

The third method for sampling hidden populations is chain-referral sampling, the best-known form of which is snowball sampling (Goodman 1961). This has traditionally been considered a form of convenience sampling about which no claims of representativeness can be made. In a now-classic article, Erickson (1979:299) argued that the sample begins with a bias because when sampling a hidden population the choice of initial subjects cannot be random, and further biases of an unknown nature are added as the sample expands during subsequent waves. Subsequent to Erickson's analysis, additional biases have been identified, so sources of bias in chain-referral samples include: (1) nonrandom choice of initial subjects, that is, the choice of "seeds"; (2) volunteerism, in which more cooperative subjects agree to participate in larger numbers or masking in which less cooperative subjects are under represented; (3) differentials in recruitment, in which some groups recruit more peers into the study than others; (4) differentials in network size, because referrals occur through network links so groups with larger personal networks will be over sampled; and (5) differentials in homophily, or tendency toward in-group recruitment, because groups with greater homophily will be over sampled. Because of these problems, chain-referral samples have traditionally been seen merely as a form of convenience sample, suitable only for pilot studies and formative research.

Despite this recognition of bias, there has been a resurgence of interest in chain-referral methods because of their unique ability to reach those who would be missed by other methods, including those who shun public gatherings and institutional affiliations. Research on the "small world problem" suggests that any two people in the country are connected by no more than six network links, the now-famous "six degrees of separation." The implication is that everyone could be reached by a maximally expansive chain-referral sample after only a handful of waves.

A prerequisite for the use of chain-referral samples to study hidden populations is that the population be linked by a "contact pattern." That is, members of the population must know one another. These contact patterns are robust in the populations upon which AIDS-prevention research has focused. Injectors form contacts when they buy drugs, and these are strengthened because regular users cultivate multiple sources to ensure continuity of supply. These bonds are further strengthened because drugs are often purchased jointly and shared. The robustness of its contact patterns makes this population ideally suited to chain-referral sampling. Therefore, the best sources of information about injectors in any community are the injectors.

Respondent-Driven Sampling

The design principle of respondent-driven sampling (RDS) is simple. If the biases associated with chain-referral methods are understood, it is possible to redesign the sampling process to eliminate those biases that are not inherent in the method, and to quantify and control those that are inherent in the method. Therefore, RDS includes both a specific method for structuring the chain-referral process to reduce one set of biases, and analytic procedures to weight the sample to compensate for others. In this way, chain-referral sampling can be made into a statistically valid sampling method.

The first source of bias is due to the selection of initial subjects. Figure 4.3 depicts the recruitment tree generated by RDS beginning with a single seed. Over the course of many waves, the sample expanded to include more than one hundred recruits.

An examination of recruitment patterns by ethnicity confirms that the choice of initial subjects does indeed introduce a bias into the sample (see table 4.3). Recruitment reflects *homophily*, a tendency to recruit persons like oneself. For
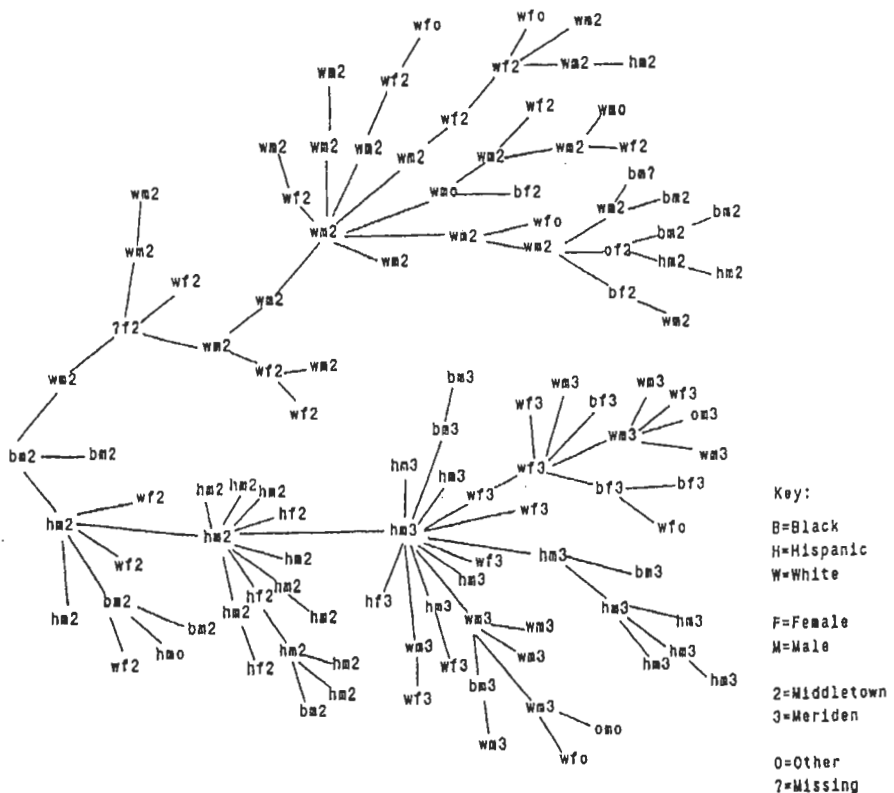

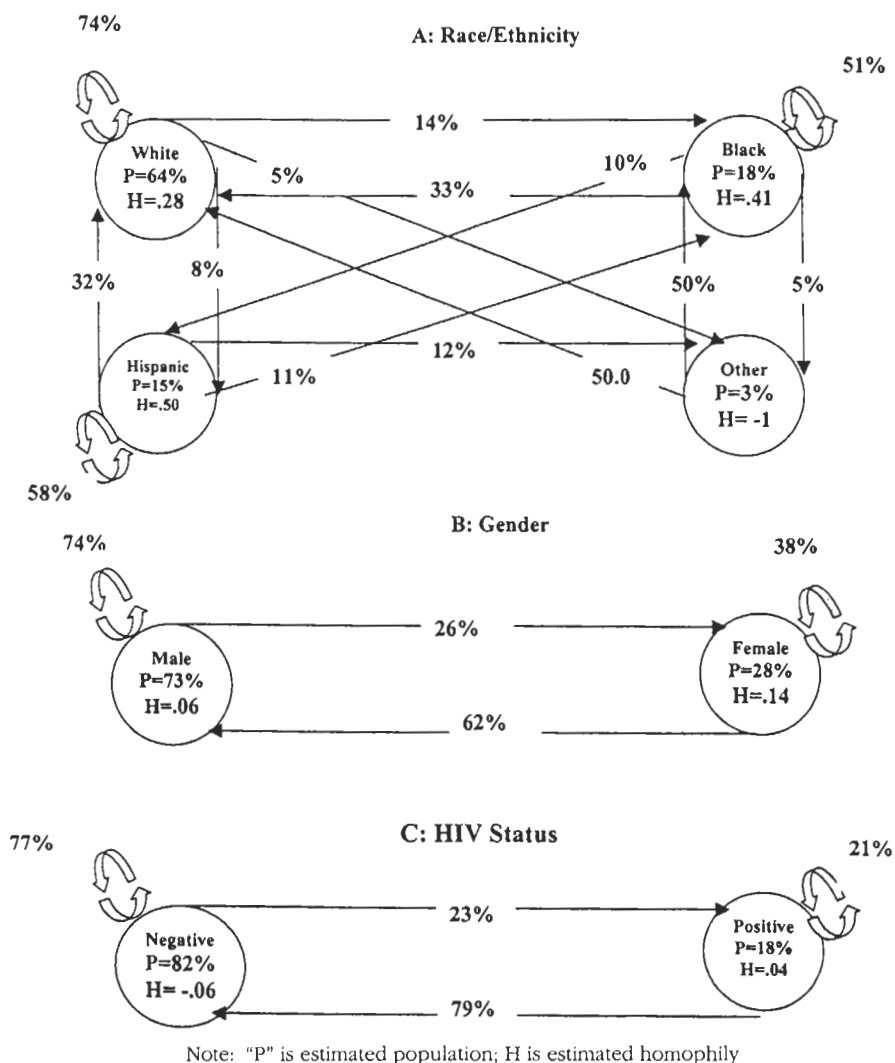**Figure 4.3. Peer Recruitment Network Beginning from a Single Seed**

Table 4.3. Race and Ethnicity of Recruits, by Race and Ethnicity of Recruiter (New London)

Race/Ethnicity of Person who Recruited	Race/Ethnicity of Recruit				
	W	B	H	O	Total
Non-Hispanic White (W) (n = 65)	73.9%	13.9%	7.7%	4.6%	100%
Non-Hispanic Black (B) (n = 39)	33.3%	51.3%	10.3%	5.1%	100%
Hispanic (H) (n = 19)	31.6%	10.5%	57.9%	0%	100%
Other (O) (n = 7)	28.6%	42.9%	28.6%	0%	100%

example, in a study of injectors in New London, Connecticut, non-Hispanic white injectors recruited, on average, 74 percent other non-Hispanic whites; Hispanics recruited 58 percent Hispanics; and non-Hispanic blacks recruited 51 percent blacks. Only the very small group in the Other category failed to recruit differentially from within.

It might seem that homophily would make chain-referral samples irrevocably biased. For example, a group that had been over represented among the seeds with which recruitment began might seem as though it would remain over represented in the sample. However, as noted above, the manner in which homophily affects recruitment as the chain-referral sample expands from wave to wave can be identified by modeling the process as a form of stochastic model known as a Markov chain (see figure 4.4). A Markov chain consists of a set of two or more states (e.g., subject characteristics such as gender or ethnicity), and transition probabilities from state to state (i.e., probabilities that a subject with a given set of characteristics will recruit a subject with each other possible set of characteristics). As an illustration of a Markov chain, see figure 4.4A, which depicts table 4.3's data on recruitment by race and ethnicity. The four states correspond to the recruiter's race and ethnicity (i.e., Hispanic, non-Hispanic black or white, and other), and the double line arrows depict the transition probabilities within and single line arrows depict transition probabilities across states. Recruitment is a stochastic process and can be visualized as a point whose location corresponds to the state of the most recent recruit, cross-state recruitment moves the point to a different state by following the arrows, and within-state recruitment keeps the point at the same location.

The conclusion from modeling the recruitment process as a Markov chain is that biases introduced by the selection of initial respondents are progressively *weakened* with each recruitment wave. The manner in which this occurs is illustrated in figure 4.5, which depicts the results of two simulations showing how the composition of each wave would have changed had recruitment begun from either Hispanic injectors (figure 4.5A) or non-Hispanic white injectors (figure 4.5B), based on projections from figure 4.4A's recruitment patterns. The vertical axes represents the percentage of injectors of each type, and the
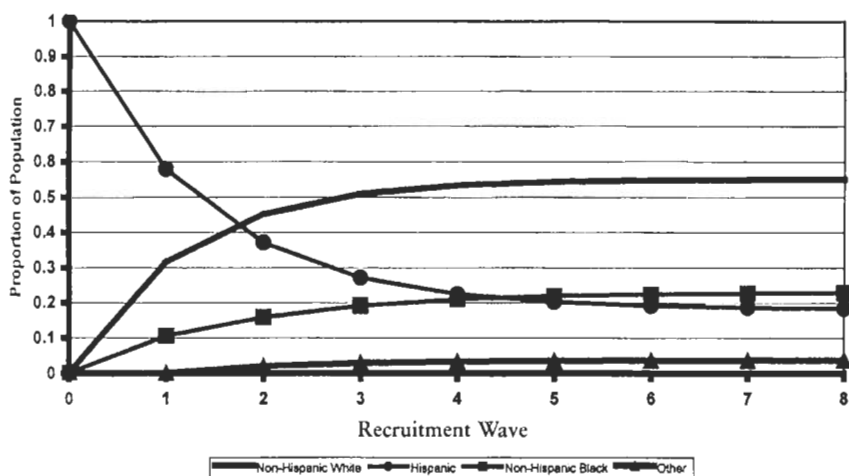
A: Race/Ethnicity

74%

51%

14%

White
P=64%
H=.28

5%

Black
P=18%
H=.41

10%

33%

32%

8%

50%

5%

Hispanic
P=15%
H=.50

11%

12%

50.0

Other
P=3%
H= -1

58%

B: Gender

74%

38%

26%

Male
P=73%
H=.06

Female
P=28%
H=.14

62%

C: HIV Status

77%

21%

23%

Negative
P=82%
H= -.06

Positive
P=18%
H=.04

79%

Note: "P" is estimated population; H is estimated homophily

Figure 4.4. Network Structures of Drug Injectors—Patterns of Association Show Varying Degrees of Homophily (New Longdon, CT)

horizontal axes represent the number of recruitment waves, where wave 0 refers to the seed or seeds, which in this exercise were assumed to be ethnically homogeneous. Wave 1 refers to the seeds' recruits; wave 2 refers the recruits' recruits, and so forth. Had recruitment begun with only Hispanic seeds the percentage of Hispanics in each wave decreases from the initial value of 100 percent, to 58 percent in the first wave, 37 percent in the second wave, eventually stabilizing at 18 percent. This stable point is termed the *equilib-*

rium, because it does not change with later waves. When equilibrium is reached, the composition of that and each additional wave is 22.9 percent non-Hispanic blacks, 55.2 percent non-Hispanic whites, 18.2 percent Hispanics, and 3.7 percent other.

A: Starting Point = All Hispanic Seeds

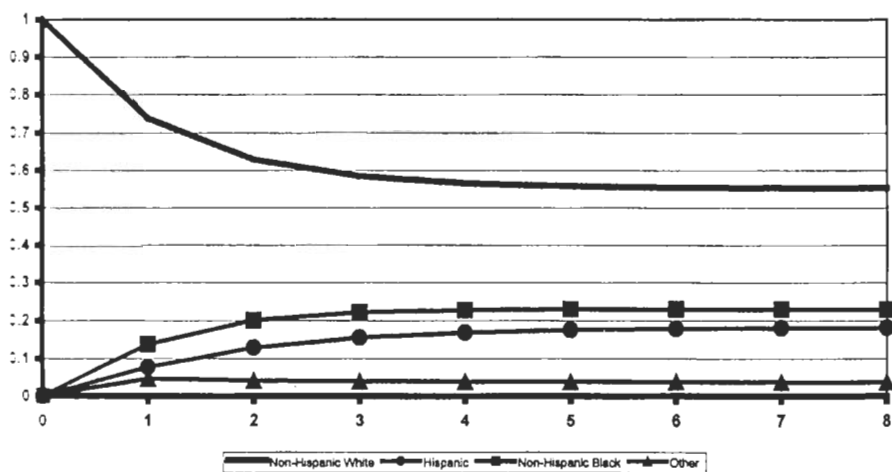
B: Starting Point = All White Seeds

Figure 4.5.   Two Simulations of Recruitment in a Respondent-Driven Sample—Race and ethnicity of recruits in a respondent-driven sample, beginning with all Hispanic or non-Hispanic white seeds.

In contrast, in the simulation where recruitment began with only non-Hispanic white seeds (figure 4.5B), the percentage of Hispanics in each wave increases, from the initial value of 0 percent, to 8 percent in wave 1, 13 percent in wave 2, and stabilizes at 18 percent in wave 5 and subsequent waves. Note that after equilibrium is attained, the composition by wave in figure 4.5B is the same as in figure 4.5A. This convergence reflects an important characteristic of RDS. If sampling is allowed to proceed through a minimum number of waves, it will attain an equilibrium that is *independent of the characteristics of the respondents from which sampling began* (Heckathorn 1997). Thus, it does not matter whether all seeds were drawn from the same group or from any mix of groups, the ultimate composition of the sample will be the same. Hence, whatever bias was introduced by the selection of initial respondents is eliminated if sampling is continued through enough waves. After the desired number of subjects is recruited, computations can be performed (see Heckathorn 1997:186) to confirm that the composition of the sample converged with the equilibrium sample composition.

To ensure that referral chains will be lengthy, respondents receive modest financial rewards for their recruiting efforts. A second means used to lengthen referral chains is a quota on recruitment rights, a limit of three recruits per respondent.

A second form of bias arises when especially cooperative respondents volunteer for recruitment in disproportionate numbers. In previous applications of RDS, several measures were employed to reduce this bias. The first was a combination of material and social incentives. Subjects were paid for the interview. A second and more consequential incentive was social, the exercise of influence by the peer recruiter. Using this dual system, even respondents for whom the material reward from the interview was irrelevant could be induced to participate through the social influence of the peer recruiter. In addition to harnessing peer pressure, another means for reducing this source of bias was to ensure that the interview site was located conveniently and in neutral turf. Otherwise, subjects would be under represented for whom the location was accessible only with difficulty, or for whom the location was threatening. Third, the interview staff was trained to treat all subjects respectfully and nonjudgmentally. In addition, amenities such as free coffee were provided. This is crucial for a sampling method that relies on peer recruitment; subjects cannot be expected to recruit peers unless their own experience was positive. To assess this potential source of bias, subjects were asked about the composition of their networks, this was compared to their recruitment patterns, and no significant discrepancies were found (Heckathorn 2002).

Thus, the use of material and social incentives to produce long referral chains serves to reduce the first two sources of bias, those due to the choice of initial subjects and volunteerism. The other three biases associated with chain-referral methods are inherent in the manner in which subjects are recruited into the sample. However, this does not mean that chain-referral sampling need be either

invalid or unreliable. For based on a sufficiently detailed understanding of these sources of bias, the sample can be weighted to compensate exactly for them, thereby producing an unbiased sample. The procedure for compensating for each of the three source of bias relies on the same logic. Information gathered during the recruitment process provides the means for quantitatively measuring this bias, and then controlling it through a weighting process.

The third source of bias derives from differential recruitment. This bias occurs when a group recruits especially effectively, because its distinctive recruitment pattern is thereby over represented in the sample—for example, if this group tends to recruit more of another group's members, the latter will be over represented in the sample. The recruitment quotas reduce this form of potential bias, but given that not all subjects fulfill their quotas, variation remains. Previous applications of RDS provide evidence for recruitment differentials. For example, in a study of injectors in Meriden, Connecticut, it was found that HIV positive respondents recruited 69 percent more than HIV negative respondents. Of course, this greater involvement of HIV positive subjects in an AIDS-prevention intervention is easily understandable. Similarly, in that town, injectors aged 26 and more recruited 20 percent more than did injectors aged 18–25, so the older injectors were more energetic. RDS weights the sample to compensate for these differentials, by mathematically projecting what the sample composition would have been had all groups recruited with equal effectiveness (Heckathorn 2002).

Biases due to differentials in network size are controlled in a logically similar manner. When network sizes are unequal, the better-connected group is over sampled. For example, in the study of several small cities in Connecticut, New London, Middletown, and Meriden, HIV positive injectors were found to have consistently larger personal networks by amounts ranging from 28 percent to 36 percent. Therefore, they were reachable by a larger number of potential recruitment chains, and were hence recruited in somewhat greater numbers. The procedure for compensating for differentials in network size is based on the recognition that in RDS, the relationship between recruiter and recruit is almost always reciprocal, where this means that a tie in one direction implies a tie in the opposite direction (Heckathorn forthcoming). For example, if A is a friend to B, then B is typically also a friend to A. In previous applications of RDS, the relationships between recruit and recruiter involved some form of ongoing personal relationship in 97 percent to 99 percent of cases depending on the site. For example, in New London, the recruiter was most commonly a friend (64 percent), followed by an acquaintance (19 percent), or spouse or other sex partner (9 percent), and only 3 percent reported having been recruited by a stranger. The presence of preexisting social relationships between recruit and recruiter results from the combination of a quota on recruitment coupons and the rewards for recruiting. If recruitment coupons were available in unlimited numbers, impersonal forms of recruiting would be possible, for example, placing piles of coupons in locations were potential recruits might gather. However,

given that coupons are given out only in very modest numbers, the overwhelming majority of respondents recruit persons with whom they already have a relationship rather than approaching strangers. The RDS method weights the sample to compensate for these differentials, by mathematically projecting what the sample composition would have been had all groups had equal network sizes (Heckathorn 2002). Similarly, biases due to differentials in homophily are controlled through mathematically projecting what the sample composition would have been had all groups had equal homophily. Finally, standard errors are computed based on a variant of boot-strapping (Heckathorn 2002).

Using RDS to Study Social Structure

An advantage of RDS over standard probability sampling is that it provides information not only about the respondents, but also about their social structure. For when respondents recruit their peers they provide information about the social network in which they are embedded, where, as in the works of Simmel, social structure is defined in terms of the structure of affiliations. Furthermore, this information is based not on self-reports, as in network sampling (Granovetter 1976), in which respondents are asked to report on their peers. In contrast in RDS, information on network links is behavioral, having been derived from recruitment records. This is significant, not only because of issues of validity and reliability, but also because analyses need not be limited to information respondents have regarding their peers. Consider, for example, figure 4.4C's depiction of structure by HIV status, which shows that HIV positive respondents are thoroughly integrated in the injector community. That is, both HIV positive and negative respondents have near zero homophily. Such an analysis would not have been possible based only on self-reports, because many injectors do not know the HIV status of their peers, and even if they did have that information, its disclosure without the permission of the peer would be a violation of the Connecticut HIV Confidentiality Law. Such problems do not arise for RDS, because respondents need only provide information regarding themselves, and network links are established behaviorally.

This approach to studying social structure is consistent with social network-based quantitative definitions (Blau 1977; Rapoport 1979). A system is said to lack structure if social relationships are formed randomly. In that case, individuals are indifferent between ties formed within and outside the group. Therefore, the proportion of within-group ties equals the proportional size of the group. As thus defined, structure can take either of two basic forms. First, *homophily* refers to a tendency to form within-group ties. Homophily in the formation of friendships was recognized before the turn of the century by Galton, and it has been found based on age, education, prestige, social class, and race and ethnicity (McPherson and Smith-Lovin 1987). Alternatively, *heterophily* or equivalently, *negative homophily*, refers to a tendency to form out-group ties, for example, tribes with exogamous marriage systems require that marriages occur outside one's clan. As thus defined, a system is *structured* if it reflects

either homophily or heterophily. Thus, homophily and heterophily are the elements out of which social structures are built. Figure 4.4 provides graphic depictions of the social structure of injectors in New London, Connecticut, with respect to race/ethnicity, gender, and HIV status.

As thus defined, structure depends both on the form of relationship considered and the type of group. If a basis for group identification is socially irrelevant, it does not serve as a basis for structural differentiation, and homophily is zero. For example, whether one is born in an odd or even month is socially irrelevant, so homophily is zero based on this status. Similarly, in the United States, blood type is socially irrelevant, so homophily is presumably zero. In contrast, in Japan where blood type is widely believed to determine interpersonal compatibility, if that belief comes to shape the formation of relationships, structure will emerge. In contrast, race and ethnicity, and other basic demographic variables, are strongly structuring. Thus, determining whether an attribute affects homophily is a way of determining its social significance. This approach to the analysis of social structure is now being employed to study jazz musicians in four cities, New York, New Orleans, Detroit, and San Francisco (Heckathorn and Jeffri 2001). The aims include determining the extent to which affiliation patterns are determined by musical form (e.g., fusion, classical, or contemporary jazz), rather than the demographic factors, such as race and class, around which affiliation patterns are structured in the larger society.

CONCLUSION

This brief summary of an ongoing research program offers a couple of lessons. One is that the stark separation traditionally existing in sociology between theoretic and applied work need not exist. James Coleman (1990) emphasized the extent to which modern societies are increasingly the product of purposive design. Most of us spend our lives embedded in markets and organizations whose guiding principles are defined by laws and regulations. Politicians play an especially important role in institutional design, at levels ranging from the halls of congress to local zoning boards, and from heads of interest groups to community activists. Decisions made within firms are no less important, particularly because, under the impact of globalization, some firms dwarf in economic power all but the largest countries. Therefore, fundamental institutions are being designed and redesigned, frequently without any systematic assessment of the consequences of those decisions. For example, in the case of AIDS prevention, the traditional approach was not adopted after an exhaustive assessment of available alternatives; it merely borrowed a design from the antipoverty programs of the 1960s. Institutional design is a process to which sociologists should be able to contribute. Indeed, after having made a profession of studying social phenomena, it would bode ill for the future of the profession if sociologists had no useful advice to offer. Many prominent sociologists have combined theoretic and applied work (e.g., James Coleman and Peter Rossi), however this remains

the exception. As the example of AIDS-prevention intervention illustrates, even apparently highly abstract theory can have important practical implications.

A second lesson from the research program concerns the inherently uncertain nature of the research enterprise. It is inherent in research that one does not know what the result will be. Otherwise, one would be an engineer. When conducing research, one must remain open to unanticipated results, whatever they may be. This involves not only the obvious requirement that one act with integrity, by accepting disappointing results when a favored theory fails an empirical test. Equally important, it involves remaining alert to opportunities to expand the project, even into wholly unanticipated directions. For example, when I began the study of norm emergence, I had no idea this would become relevant to AIDS prevention; and when beginning the design of the AIDS-prevention intervention, I had no idea that this would lead to improved means for sampling hidden populations and studying social structure. This experience is by no means atypical, even in other research programs in which I have been involved. For example, a study of bargaining resulted in my development of a formal model, resistance theory (Heckathorn 1980,1983), that serves as both an element of network exchange theory (Markovsky, Willer, and Patton 1988), and also provided the basis for a transaction resource theory that was used to analyze negotiation during the U.S. Constitutional Convention of 1787 (Heckathorn and Maser 1987a), the history of regulation of business in the United States (Heckathorn and Maser 1987b), disclosure rules and default provisions in contract law (Coleman, Heckathorn, and Maser 1989), congressional decision making (Heckathorn and Maser 1990), and the debate from 1787 to 1789 over ratification of the U.S. Constitution (Anthony, Heckathorn and Maser 1994). As a researcher, one must remain alert to unanticipated opportunities to move the project into a new direction. Wherever the path leads, one must be prepared to follow.

NOTE

REFERENCES

Anthony, Denise L., Douglas D. Heckathorn, and Steven M. Maser. 1994. "Rational Rhetoric in Politics: The Debate over Ratifying the U.S. Constitution." *Rationality and Society* 6:489–518.

Blau, Peter M. 1977. *Inequality and Heterogeneity*. New York: Free Press.

Broadhead, Robert S., and Kathryn J. Fox. 1990. "Takin' It to the Streets: AIDS Outreach as Ethnography." *Journal of Contemporary Ethnography* 19:322–48.

Broadhead, Robert S., and Douglas D. Heckathorn. 1994. "AIDS Prevention Outreach among Injection Drug Users: Agency Problems and New Approaches." *Social Problems* 41:473–95.

Broadhead, Robert S., Douglas D. Heckathorn, Frederick Altice, Yael van Hulst, Michael Carbone, Gerald Friedland, Patrick O'Connor, and Peter Selwyn. Forthcoming. "Increasing Drug User's' Adherence to HIV Theropeutics." *Social Science and Medicine*.

Broadhead, Robert S., Douglas D. Heckathorn, Jean-Paul C. Grund, L. Synn Stern, and Denise L. Anthony. 1995. "Drug Users versus Outreach Workers in Combating AIDS: Preliminary Results of a Peer-Driven Intervention." *Journal of Drug Issues* 25:531–64.

Broadhead, Robert S., Douglas D. Heckathorn, David Weakliem, Denise Anthony, Heather Madray, Robert Mills, and James Hughes. 1998. "Harnessing Peer Networks as an Instrument for AIDS Prevention: Results from a Peer-Driven Intervention." *Public Health Reports* 113, supplement 1:42–57.

Brown, B., and G. M. Beschner. 1993. *Handbook on Risk of AIDS: Injection Drug Users and Sexual Partners.* Westport, CT: Greenwood Press.

Brown, Lawrence D., Morris L. Eaton, David A. Freedman, Stephen P. Klein, Richard A. Olshen, Kenneth W. Wachter, Martin T. Wells, and Donald Ylvisaker. 1999. "Statistical Controversies in Census 2000." *Technical Report 537*, Department of Statistics, U.C. Berkeley.

Coleman, James S. 1990. *Foundations of Social Theory.* Cambridge, MA: Belknap Press.

Coleman, Jules L., Douglas D. Heckathorn, and Steven M. Maser. 1989. "A Bargaining Theory Approach to Default Provisions and Disclosure Rules in Contract Law." *Harvard Journal of Law and Public Policy* 12:639–709.

Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior.* New York: Plenum.

Erickson, Bonnie H. 1979. "Some Problems of Inference from Chain Data." *Sociological Methodology* 10:276–302.

Gilham, Steven A. 1982. "The Marines Build Men: Resocialization in Recruit Training." Pp. 231–41 in Reid Luhman (ed.), *The Sociological Outlook.* Belmont, CA: Wadsworth.

Goodman, L. A. 1961. "Snowball Sampling." *Annals of Mathematical Statistics* 32:148–70.

Granovetter, Mark. 1976. "Network Sampling: Some First Steps." *American Journal of Sociology* 83:1287–1303.

Heckathorn, Douglas D. 1980. "A Unified Model for Bargaining and Conflict." *Behavioral Science* 25:261–84.

———. 1983. "Extensions to Power-Dependence Theory: The Concept of Resistance." *Social Forces* 61:1206–31.

———. 1988. "Collective Sanctions and the Emergence of Prisoner's Dilemma Norms." *American Journal of Sociology* 94:535–62.

———. 1989. "Collective Action and the Second-Order Free-Rider Problem." *Rationality and Society* 1:78–100.

———. 1990. "Collective Sanctions and Compliance Norms: A Formal Theory of Group-Mediated Social Control." *American Sociological Review* 55:366–84.

———. 1991. "Extensions of the Prisoner's Dilemma Paradigm: The Altruist's Dilemma and Group Solidarity." *Sociological Theory* 9:34–52.

———. 1993. "Collective Action and Group Heterogeneity: Voluntary Provision versus Selective Incentives." *American Sociological Review* 58:329–50.

———. 1997. "Respondent Driven Sampling: A New Approach to the Study of Hidden Populations." *Social Problems* 44:174–99.

———. 2002. "Respondent Driven Sampling II: Deriving Valid Population Estimates from Chain-Referral Samples of Hidden Populations." *Social Problems* 49:11–34.

Heckathorn, Douglas D., Robert S. Broadhead, Denise L. Anthony, and David L. Weakliem. 1999. "AIDS and Social Networks: Prevention through Network Mobilization." *Sociological Focus* 32:159–79.

Heckathorn, Douglas D., and Joan Jeffri. 2001. "Finding the Beat: Using Resondent-Driven Sampling to Study Jazz Musicians." *Poetics* 28:307–29.

Heckathorn, Douglas D., and Steven M. Maser. 1987a. "Bargaining and Constitutional Contracts." *American Journal of Political Science* 31:142–68.

———. 1987b. "Bargaining and the Sources of Transaction Costs: The Case of Government Regulation." *Journal of Law Economics and Organization* 3:69–98.

———. 1990. "The Contractual Architecture of Public Policy: A Critical Reconstruction of Lowi's Typology." *Journal of Politics* 52:1101–23.

Heggoy, Alf Andres. 1972. *Insurgency and Counterinsurgency in Algeria.* Bloomington: Indiana University Press.

Institute of Medicine. 1995. *Assessing the Social and Behavioral Science Base for HIV/AIDS Prevention and Intervention.* Washington, D.C.: National Academy Press.

Kemeny, John G., and J. Laurie Snell. 1960. *Finite Markov Chains.* Princeton, NJ: Van Nostrand.

Laumann, Edward O., John H. Gagnon, Stuard Michaels, Robert T. Michael, and James S. Coleman. 1989. "Monitoring the AIDS Epidemic in the United States: A Network Approach." *Science* 244:1186–89.

Markovsky, Barry, David Willer, and Travis Patton. 1988. "Power Relations in Exchange Networks." *American Sociological Review* 53:220–36.

McPherson, J. Miller, and Lynn Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* 52:370–79.

Needle, Richard H., and Susan Coyle. 1997. "Community-Based Outreach Risk Reduction Strategy to Prevent HIV Risk Behaviors in Out-of-Treatment Injection Drug Users (IDUs)." Paper presented at the National Institutes of Health Consensus Development Conference on Interventions to Prevent HIV Risk Behaviors, Bethesda, MD, Feb. 11–13.

Oliver, Pamela E., Gerald Marwell, and Ruy Teixeira. 1985. "A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action." *American Journal of Sociology* 91:522–56.

Olson, Mancur. 1965. *The Logic of Collective Action.* Cambridge: Harvard University Press.

Preble, Ed, and J. J. Casey. 1969. "Taking Care of Business—The Heroin User's Life on the Street." *International Journal of Addictions* 4:1–24.

Rapoport, Anatol. 1979. "A Probabilistic Approach to Networks." *Social Networks* 2:1–18.

Rapoport, David C. 1984. "Fear and Trembling: Terrorism in Three Religious Traditions." *American Political Science Review* 78:658–77.

Sergeyev, Boris, Tatyana Oparina, Tatyana P. Rumyantseva, Valerii Volkanevskii, Robert S. Broadhead, Douglas Heckathorn, and Heather Madray. 1999. "HIV Prevention in Yaroslavl, Russia: A Peer-Driven Intervention and Needle Exchange." *Journal of Drug Issues* 29:777–804.

Sudman, Seymour, and Graham Kalton. 1986. "New Developments in the Sampling of Special Populations." *Annual Review of Sociology* 12:401–29.

Watters, John K., and Patrick Biernacki. 1989. "Targeted Sampling: Options for the Study of Hidden Populations." *Social Problems* 36:416–30.