AIDS and Behavior, Vol. 6, No. 1, March 2002 (© 2002)

Extensions of Respondent-Driven Sampling: A New Approach to the Study of Injection Drug Users Aged 18–25

Douglas D. Heckathorn,[1,5] Salaam Semaan,[2] Robert S. Broadhead,[3] and James J. Hughes[4]

Received June 22, 2000; revised June 8, 2001; accepted July 18, 2001

Researchers generally use nonprobability methods such as chain-referral sampling to study populations for which no sampling frame exists. Respondent-driven sampling is a new form of chain-referral sampling that was designed to reduce several sources of bias associated with this method, including those from the choice of initial participants, volunteerism, and masking. This study expands this method by introducing "steering incentives," supplemental rewards for referral of members of a specific group, injection drug users (IDUs) aged 18–25. The results are based on an interrupted time series analysis in which 196 IDUs from Meriden, CT, were interviewed before introduction of the steering incentives, and another 190 were interviewed afterwards. The steering incentives increased the percentage of younger IDUs sampled by 70%. We compared recruitment patterns with institutional data and self-reported personal networks to determine representativeness and whether volunteerism or masking were present. The results indicated that steering incentives helped to increase recruitment of younger IDUs, that the sample was representative, and that both volunteerism and masking were modest.

KEY WORDS: injection drug user; sampling methodology; hidden population; AIDS prevention; HIV.

INTRODUCTION

Injection drug users (IDUs) aged 18–25 years are at risk for HIV infection, yet the number of these "younger" IDUs recruited for HIV risk-reduction studies and interventions have been small (Coyle *et al.*, 1998). Traditional probability sampling methods, such as household surveys, are unsuitable for reaching younger IDUs and other hidden populations, because response rates are low and responses lack candor (Spreen and Zwaagstra, 1994). Similarly, the street-based location sampling methods that have dominated much risk-reduction research (Semaan *et al.*, 1998), such as targeted sampling (Watters and Biernecki, 1989), tend to recruit IDUs who spend considerable time on the street, especially older male IDUs. In contrast, chain-referral methods sample based on network connections among respondents, and so even respondents who spend little time on the street can be reached. However, use of this method in a study of IDUs in New London, CT, produced only 7% younger IDUs (Heckathorn *et al.*, 1999).

We used *respondent-driven sampling* (RDS), a form of chain-referral network sampling designed to overcome many of the problems generally attributed to chain-referral sampling (Heckathorn, 1997). In previous work, RDS also served as the recruitment mechanism for a *peer-driven HIV-prevention intervention* (Broadhead *et al.*, 1998; Broadhead and Heckathorn, 1994; Heckathorn *et al.*, 1999). In this paper, we extend the RDS method by introducing *steering incentives* that provide respondents and their peer recruiters with bonuses for recruiting members of a specific group, "younger" IDUs, 18–25 years old.

[1]Department of Sociology, Cornell University, Ithaca, New York.
[2]Behavioral Intervention Research Branch, Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia.
[3]Department of Sociology, University of Connecticut, Storrs, Connecticut.
[4]Office of Institutional Research, Eastern Connecticut State University, Windham, Connecticut.
[5]Correspondence should be directed to Douglas Heckathorn, PhD, Department of Sociology, Cornell University, Ithaca, New York 14853; e-mail: douglas.heckathorn@cornell.edu.

1090-7165/02/0300-0055/0 © 2002 Plenum Publishing Corporation

According to Spreen and Zwaagstra (1994), the best that one can expect from a sample of a hidden population is a broad cross-section of the hidden population. Samples that satisfy this modest goal are considered *representative*. Our aims here are, first, to assess the effectiveness of these incentives for increasing recruitment of younger IDUs, by comparing the composition of the sample before introduction of the steering incentive with the composition of the sample after introduction of the incentive. Second, we assess the correspondence between the sample and the composition of the hidden population as measured using three types of institutional data and census data. Finally, we test for volunteerism and masking—overrecruitment of especially cooperative participants or under recruitment of less cooperative participants, respectively—by comparing the composition of the sample and the composition of self-reported personal networks. We begin by discussing the importance of developing means for sampling younger IDUs and review the techniques and advantages of RDS.

Sampling Younger IDUs

To date, more than one third of all reported AIDS cases in adults in the United States has been in IDUs, their heterosexual partners, and children of mothers who are IDUs or sexual partners of IDUs (Centers for Disease Control and Prevention [CDC], 1997). The proportion of AIDS cases among younger IDUs is substantial. For men aged 13–19 years, and 20–24 years, injection drug use accounts for 6 and 12% of AIDS cases, respectively. For women in these two age groups, the respective proportions are 15 and 29% (CDC, 1997). More than half of all new HIV infections in the United States are estimated to occur in IDUs, their sex partners, and their children (Holmberg, 1996). According to data from states that have integrated HIV and AIDS reporting, from January 1994 through June 1997, 14% of HIV cases occurred in persons aged 13–24 years, and 449 (6%) of these cases were due to injection drug use (*MMWR*, 1998).

Part of the difficulty of sampling younger IDUs stems from the fact that they constitute a "hidden population," as do other IDUs. That is, there exists no list of population members to serve as a sampling frame.

Targeted sampling has been used in many HIV risk-reduction interventions. It involves two basic steps. First, field researchers are deployed to map a target population. To the extent that field researchers succeed in penetrating the local networks linking potential respondents, this prevents the undersampling that traditional approaches would produce. Second, field researchers are deployed at diverse sites identified by the ethnographic mapping to recruit a prespecified number of participants. This ensures that participants from different areas and subgroups will appear in the final sample. Using this approach, a high proportion of older are recruited (Braunstein, 1993; Liebman and Thomas, 1992; Robles *et al*., 1993; Watters, 1993; Watters and Biernecki, 1989), for several reasons. First, targeted sampling relies on outreach workers to venture into communities where the prevalence of drug use is high; these workers generally work during daylight hours. Outreach workers also tend to restrict their recruitment efforts to highly visible public areas and to situate themselves so that IDUs will know where to find them (Broadhead and Fox, 1993; Johnson *et al*., 1990; Rivera-Beckman, 1992). As a result, targeted sampling tends to undersample younger IDUs who tend to spend less time on the streets. However, were younger IDUs to be specifically targeted, they might be enrolled in larger numbers, but younger IDUs recruited from the streets may not prove representative of the population of younger IDUs.

Finding better ways to recruit representative samples of such hidden populations would result in important benefits: (1) the full spectrum of at-risk persons could be recruited for prevention interventions, (2) research results could be generalized to the populations from which the samples were drawn, and (3) appropriate policies could be developed. Furthermore, developing of means to oversample groups of special interest would provide the means to draw stratified samples.

Respondent-Driven Sampling

Recognition of the problems inherent in street-based recruitment has led to an increased focus on chain-referral methods. These methods are predicated on the recognition that peers are better able than field workers to reach other members of the hidden population. Sampling begins with a set of initial participants who serves as "seeds," and expands in waves, where Wave 1 consists of participants referred by the seeds, Wave 2 consists of participants referred by the first-wave participants, and so each recruitment is a link in the recruitment chain. These methods have been widely used to study hidden populations, including drug users (Frank and Snijders, 1994;

Klovdahl, 1989; Spreen and Zwaagstra, 1994; Sudman *et al*., 1988).

A number of distinct chain-referral sampling methods have been developed, which vary depending on the manner in which referrals take place, the number of referrals per respondent, the maximum number of waves, and the type of data gathered. For example, in Klovdahl's "random walk" approach, each respondent lists persons whom they know who fall within the target population, along with their contact, information; the researcher then randomly selects one person for contact, up to a maximum of three waves. Therefore, each initial participant yields a maximum of three other participants. In contrast, Frank and Snijders (1994) recommend starting with a large and diverse number of initial participants and conducting only a single wave.

Sampling bias in chain-referral sampling has been extensively analyzed (Erickson, 1979; Heckathorn, 1997). RDS is a form of chain-referral sampling that was designed to eliminate sources of bias that are not inherent in the method (i.e., bias because of the selection of initial respondents), and to reduce other sources of bias (i.e., biases because of volunteerism or masking).

First, RDS incorporates a method for removing sampling bias introduced by the (typically arbitrary) selection of respondents from which referrals begin. Erickson (1979) argued that the selection of initial respondents introduces an unknown bias into chain-referral samples that is further compounded with each additional recruitment wave. This occurs because recruitment reflects affiliation patterns, and as has been recognized since Galton's studies more than a century ago, friendships and other forms of affiliation tend to occur among persons who are similar in levels of education, income, ethnicity, and interests (McPherson and Smith-Lovin, 1987). This tendency toward affiliation with persons who are similar is termed *homophily*. In the absence of homophily, recruitment from the in-group would reflect merely the group's size, and so in-group recruitment would sum to 100%. However, as Fig. 1's depiction of recruitment relationships illustrates, recruitment reflects substantial homophily in that for each of the three categories examined, race/ethnicity, gender, and drug preference, in-group recruitment exceeds 100%. Homophily is greatest for ethnicity, where in-group recruitments sum to 162%. That is, non-Hispanic Whites recruited other non-Hispanic Whites 81% of the time, and the corresponding figures for Hispanics was 45%, and for non-Hispanic Blacks, 36%. Only the small number of participants in the "other" category failed to recruit from within. Homophily by gender was weaker, with in-group recruitment summing to 116%; and for drug preference homophily is weak, with in-group recruitment summing to only 105%.

Homophily is a potential source of bias in chain-referral samples, because recruitment reflects affiliation patterns, that is, homophily. It might therefore seem that the composition of the sample would merely reflect the characteristics of the seeds with which it began. However, Heckathorn (1997) showed that as recruitment chains grow progressively, as the sample expands from wave to wave, this bias is progressively weakened. As the sample grows in size from wave to wave, the composition of the sample ceases to change, and the stable sample composition that is attained is termed the *equilibrium* sample composition. The implication is that it does not matter whether all seeds were drawn from one group or diverse groups, the ultimate composition of the sample will be the same, so long as the number of waves is sufficiently large.

The manner in which this occurs is illustrated in Fig. 2, which depicts the results of two simulations showing how the composition of each wave would have changed had recruitment begun from either one or more Hispanic IDUs (Fig. 2(A)) or one or more non-Hispanic White IDUs (Fig. 2(B)), based on projections from Fig. 1(A)'s recruitment patterns. The vertical axis represents the percentage of IDUs of each type, and the 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 would have decreased from the initial value of 100%, to 45% in the first wave, 27% in the second wave, eventually stabilizing at 17% after five waves. This equilibrium does not change with later waves.

In contrast, in the simulation where recruitment began with only non-Hispanic White seeds (Fig. 2(B)), the percentage of Hispanics in each wave increases, from the initial value of 0%, to 14% in Wave 1, 16% in Wave 2, and stabilizes at 17% in Wave 4 and subsequent waves. Note that after equilibrium is attained, the composition by wave in Fig. 2(B) is the same as in Fig. 2(A). 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*

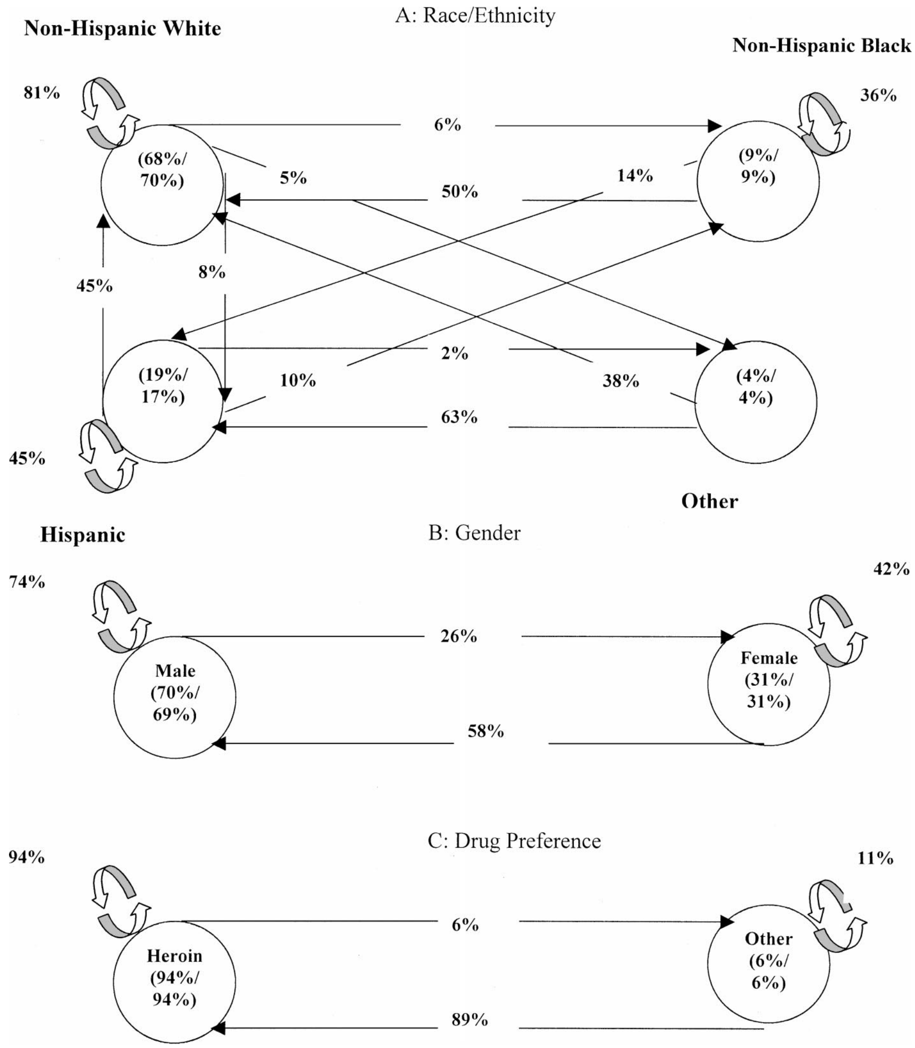

Fig. 1. Patterns of recruitment of drug injectors show differing degrees of homophily, that is, preference for ties within the group. The numbers in each node are percentage in the preincentive sample and the percentage in the theoretically computed equilibrium, where the latter is the asymptote approached as the number of recruitment waves increases. The arrows represent probabilities of recruitment by respondent group.

began (Heckathorn, 1997). (To compute this equilibrium and the number of waves required to reach it, see the Appendix.) Thus, according to the results of this simulation, it does not matter whether all seeds were drawn from one group or the other, 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 participants is recruited, computations can be performed (see Heckathorn, 1997, p. 186), to confirm that the composition of the sample converged with the equilibrium sample composition. This shows that a chain-referral sampling method can be reliable if the number of waves is sufficiently large.

Furthermore, previous applications of RDS showed that the number of waves required for the sample to reach equilibrium is not large, generally no more than four to six. Furthermore, computational procedures were provided (see Heckathorn, 1997, p. 186) with which to confirm that the composition of the actual sample converged with the equilibrium sample composition.

On the basis of this conclusion that to be reliable a chain-referral sample must have a large number of waves, RDS employs three mechanisms to lengthen

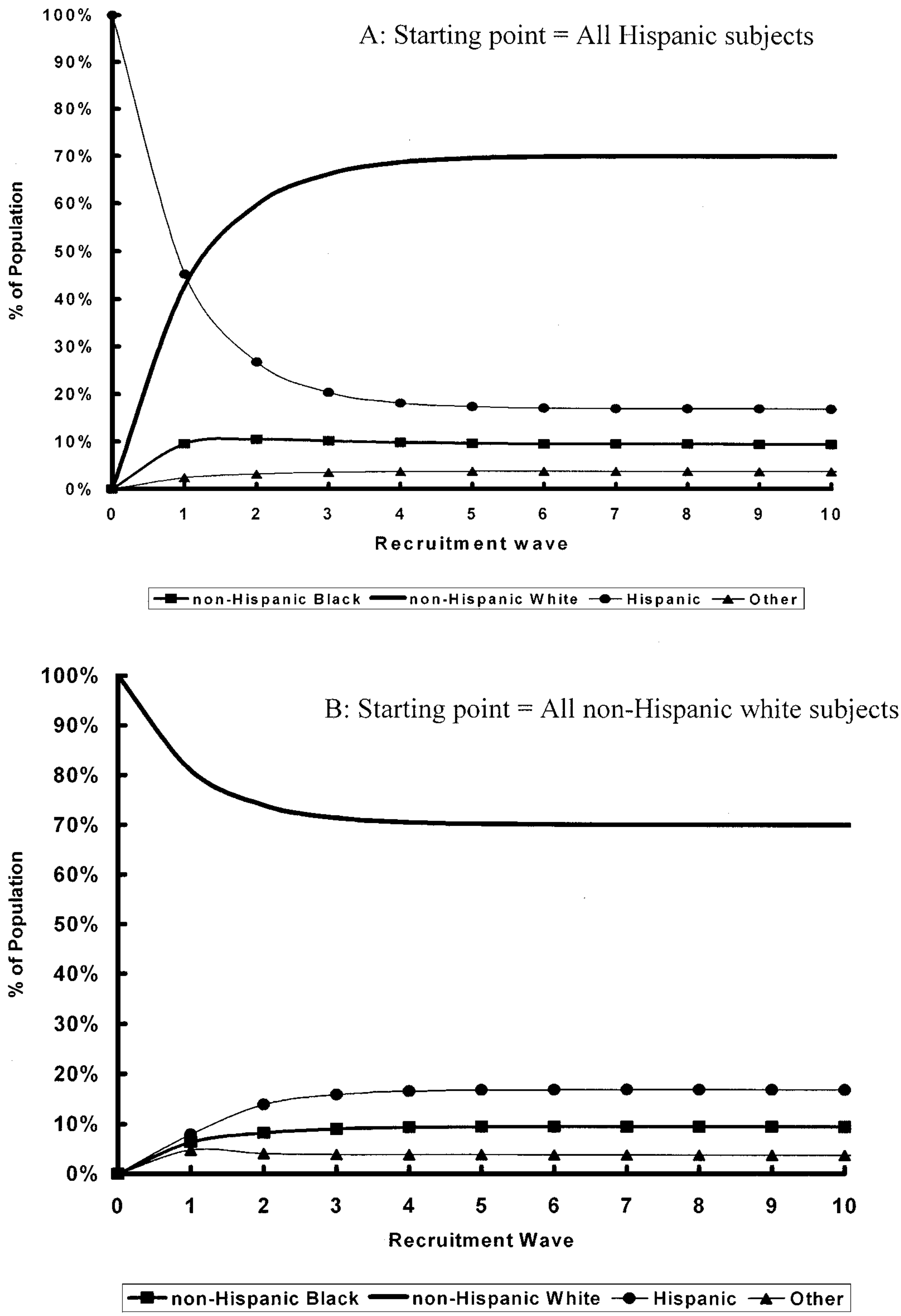

Fig. 2. Two simulations of recruitment in a respondent-driven sample: Race and ethnicity of recruits in a respondent-driven sample, beginning with all Hispanic or non-Hispanic White seeds.

recruitment chains. First, it uses a dual system of recruitment incentives that includes not only the customary payment for being interviewed, but also a payment for recruiting peers. Incentives in our RDS studies, including this one, were implemented through recruitment coupons. Respondents were told that they would receive a payment for each IDU peer to whom they gave a coupon, and who then brought that coupon to the interview. Coupons were tracked by serial numbers, and the recruitment incentives were computed by a software program developed for this project (Heckathorn *et al.*, 2001). Such software is important for two reasons. First, given that hundreds of recruitment coupons are in circulation during the course of sampling, finding out what rewards participants have earned would be a slow and tedious process in the absence of computer assistance. Second, the reward system creates incentives for respondents to seek to participate in the study multiple times, and even to seek to present themselves as their own recruit, and so the software included a participant identification procedure based on observable characteristics (e.g., gender, eye color, scars, and tattoos), and biometric measures (i.e., wrist width and forearm length).

A second means employed to lengthen recruitment chains in RDS is rationing of recruitment rights. For a given sample size, the smaller the number of recruits permitted per participant, the longer the number of waves required for the target sample size to be attained. Rationing was implemented through quotas on recruitment coupons. Each respondent was given three coupons for recruiting peers. We also rationed recruitment for another reason. Rationing prevents the emergence of semiprofessional recruiters, who might then battle over turf, and it provides all respondents with the opportunity to participate as a peer recruiter (Broadhead and Heckathorn, 1994).

Third, while in snowball and random-walk forms of chain-referral sampling, researchers recruit respondents for each wave on the basis of lists of names and contact information provided by respondents from the previous wave, in RDS respondents are recruited by peers. The former procedure is problematic because respondents frequently lack accurate contact information for their peers (Klovdahl, 1989), and so recruitment waves are terminated prematurely. The latter procedure was also adopted for another reason. When sampling a hidden and stigmatized population, recruitment by researchers requires respondents to violate the confidentiality of peers by providing their names and contact information to the researcher, and so this procedure is now forbidden by most Institutional Review Boards. In contrast, in peer recruitment as used in the RDS method, recruits themselves decide whether to become known to researchers, and so respondents do not violate one another's confidentiality.

This combination of three procedures produced very long referral chains by the standards of the chain-referral literature. Some of the respondents who served as seeds initiated recruitment chains that, over the course of more than a dozen waves, resulted in more than 100 recruits (Heckathorn *et al.*, 1999). Viewed theoretically, long referral changes are desirable for a number of reasons. First, the longer the chains the deeper the sampling process penetrates into the network structure of the hidden population, thereby reaching into even the more isolated sectors of the population. Second, given the geometrically expanding nature of the RDS recruitment process, longer referral chains ensure that the sociometric distance between the seeds and the bulk of the sample will be large, thereby further assuring the diversity and representativeness of the sample. Third, long referral chains ensure that equilibrium will be attained even if equilibrium is reached comparatively slowly.

The measures introduced to lengthen referral chains produce a recruitment process that is intended to be sufficiently robust to reduce another source of potential bias volunteerism and masking. Recruitment incentives are based on both a material component—the $20 reward for the initial interview—and a social component, the peer pressure exerted by the recruiter. Therefore, even those participants to whom the material reward was not salient may respond to peer pressure from the recruiter. This is consistent with Erickson's suggestion that these sources of bias can be reduced if one were to "build in added incentives" (Erickson, 1979, p. 229).

To ensure that the exercise of social influence by peer recruiters did not become coercive, recruitment incentives were modest, and to confirm that this measure was effective, participants were asked whether their recruiter has used excessive pressure during recruiting. Neither in the study reported here, nor in any of the other implementations of the RDS method, were any instances of coercive recruitment reported.

METHODS

A total of 386 IDUs from Meriden, CT, were interviewed. Meriden has 59,500 residents (1990 census), and substantial injection drug use. As of

June 1999, 123 cases of AIDS had been diagnosed in Meriden among adults and adolescents, 74 (60%) of which were associated with injection drug use (Carley, 1999). In addition to being interviewed, respondents were provided with HIV-prevention education, HIV-prevention materials including bleach kits and syringes, and the opportunity for HIV testing as part of a HIV-prevention intervention (Broadhead *et al.*, 1998). The intervention included an initial interview and three subsequent follow-up interviews. Participants were told that the aim of the project was prevention of HIV among active injectors.

Our study population was persons 18 years of age or older who had injected drugs during the previous 30 days. We did not require a stable address or a telephone. To ensure the recruitment of IDUs, injection drug use was determined by an eight-step screening. More than 80% of the respondents were selected at the first step—the presence of track marks—and the rest were selected at later steps involving the demonstration of detailed knowledge of injection procedures (Broadhead *et al.*, 1998). We assessed age with a screening question about birth date, assuming that most IDUs would be unable to back-calculate quickly: project interviewers believed that this screening mechanism worked well. The project operated out of storefronts in areas of known drug use. Three interviewers scheduled appointments for interviews 3–4 days a week. Interviews were conducted in English or Spanish, whichever the respondent preferred.

The study had two phases. During Phase 1, no steering incentives were used. The incentive was $10 for recruiting an IDU peer, and an additional reward of up to $10 for educating peers about HIV. As in most other intervention projects, respondents were paid for baseline interviews ($20). Phase 1 lasted from March 1995 to October 1997, during which 196 Meriden respondents completed the baseline interview.

During Phase 2, steering incentives were introduced: a $10 bonus for recruiting a younger IDU, between 18 and 25 years old, and younger respondents were paid $10 extra for a total of $30 for their baseline interview. Thus, the steering incentive doubled the reward for recruiting IDUs who fell in the younger age group, and increased by one half the reward for younger IDUs who were interviewed. Phase 2 lasted from October 1997 to June 1998, during which 190 Meriden respondents completed baseline interviews.

Ideally, in an interrupted time series design, the only difference between the conditions is the change in the experimental treatment condition, for example, here, the introduction of incentives. However, another change proved necessary. Four months after the introduction of the steering incentives (i.e., in February 1998), the interview site was moved from Middletown to Meriden. The Middletown site had appeared capable of serving both cities. The change was made because bus service between the two towns had been curtailed to twice daily (8:00 a.m. and 6:00 p.m.) instead of hourly, making it difficult for Meriden IDUs who lacked access to a car to travel the 15 miles to the Middletown interview site. Because the resulting decline in recruitment from Meriden threatened continued access to IDUs, we began conducting interviews in Meriden. During the study period, there were no other HIV prevention interventions or HIV-prevention education campaigns operating in Middletown or Meriden.

RESULTS AND DISCUSSION

Effect of the Steering Incentives

After the implementation of the steering incentives, the proportion of younger Meriden IDUs in the sample increased 70% from the preincentive baseline, that is, from 23% (45/196) to 39% (74/190), and this result is significant ($\chi^2 < .01$; Table I). Similarly, when only peer-recruited participants are considered, that is, when the seeds are excluded (see Table II), the percentage of younger IDUs increased 60%, from the preincentive baseline of 24% (45/190) to 38% (71/187). These results suggest that the steering incentives increased the number of interviews with younger IDUs.

To assess the effect of the change in interview location, we performed several analyses. First, because the change in interview site took place only 4 months after incentives were introduced, we divided the

Table I. Demographics Before and After Introduction of Steering Incentives, Meriden, CT

	Before incentives ($n = 196$)	After incentives ($n = 190$)
Non-Hispanic Black	13	9
Hispanic	22	55
Non-Hispanic White	67	33
Other	1	0
Age, 18–25	23	39*
Female	30	25

Note. Both initial participants (i.e., "seeds") and peer-recruited participants are included. Values are given in percentage.
$^*\chi^2 < .01$.

Table II. Age of Recruit and Recruiter, Before and After Introduction of Steering Incentive

	Age of recruit		
Age of recruiter	Age 26+	Age 18–25	Total
Before incentives[a]			
Age 26+	135 (83.9%)	26 (16.1%)	161 (100%)
Age 18–25	10 (34.5%)	19 (65.5%)	29 (100%)
Total distribution of recruits	145 (76.3%) 68.1%	45 (23.7%) 31.9%	190 (100%) 100%
Equilibrium	.345/(1 − .839 + .345)	.161/(1 − .655 + .161)	100%
After incentives[b]			
Age 26+	86 (69.9%)	37 (30.1%)	123 (100%)
Age 18–25	30 (46.9%)	34 (53.1%)	64 (100%)
Total distribution of recruits	116 (62.0%) 60.9%	71 (38.0%) 39.1%	187 (100%) 100%
Equilibrium	.469/(1 − .699 + .469)	.301/(1 − .531 + .301)	

Note. Initial participants (i.e., “seeds”) are excluded from this table because they did not have a recruiter.
[a]Mean Discrepancy, Distribution of Recruits and Equilibrium = 8.2%; maximum number of steps required for equilibrium be approximated within 2% = 5; and System Homophily Index = .49.
[b]Mean Discrepancy, Distribution of Recruits and Equilibrium = 1.1%; maximum number of waves required for equilibrium be approximated within 2% = 3; and System Homophily Index = .23.

postincentive Meriden sample into two subgroups, respondents interviewed before the change in location and participants interviewed afterward. From the baseline of 23% (45/196) younger Meriden IDUs interviewed before introduction of the incentives, the proportion of younger IDUs increased slightly to 29% (4/14) for the small subgroup interviewed at the original interview site, although this result did not attain statistical significance. The percentage of younger IDUs increased to 40% (70/176) for the larger subgroup interviewed at the Meriden site, and this result is highly significant ($\chi^2 < .001$). Thus, recruitment of younger IDUs increased in both subgroups as a result of steering incentives, though to differing degrees and with differing levels of significance.

Geographic Filtering Effect in a Respondent-Driven Sample

A remarkable feature of RDS is that respondents can be drawn from wide areas. When interviews were conducted in Middletown, 196 Meriden IDUs traveled the 15 miles separating these cities. Racial or ethnic differentials in access to transportation therefore constituted a potential source of sampling bias. In contrast, when the interview site was moved to Meriden, the interview site was a short walk, and so no respondent would be excluded because of limited access to means of transportation. Our results provide a fortuitous opportunity to assess this bias.

When we compared the pre- and postincentive respondents by ethnicity, the differences are substantial (see Table I). The percentage of Hispanics more than doubled and the percentage of non-Hispanic Whites decreased correspondingly. This difference seems to have resulted from ethnicity-based differentials in access to vehicles: 63% of Hispanics reported in the preincentive interview that they lacked access to a vehicle, only one third of non-Hispanics (36%) reported a lack of access.

This finding suggests a significant limitation of samples selected by using RDS. Although respondents may be drawn from considerable distances, differential access to transportation may produce a *geographic filtering effect* in which respondents who travel from areas lacking public transportation will be oversampled if they have private means of transportation. Therefore, samples selected by using RDS are most representative when respondents are local (i.e., when proximity to the interview site is such that transportation is not a problem). Samples are less representative when respondents are not local residents. For example, what had seemed to be a predominantly non-Hispanic White drug use area turned out to be predominantly Hispanic when barriers to transportation to the interview site were removed by relocating the interview site. For this reason, residents and nonresidents should be analyzed separately, as we have done (Heckathorn, 1997).

Finally, the influence of the geographic filtering effect on the incentives was assessed. Ethnicity, not age, was strongly linked to access to automobiles. In the preincentive interview, the proportions of younger and older IDUs who reported that they lacked access to vehicles were similar, 44 and 39%, respectively, a difference that is not statistically significant. It is therefore reasonable to conclude that the geographic filtering effect did not significantly influence the assessment of the impact of the steering incentives.

Differential Responses by Age to the Steering Incentives

Comparing the recruitment patterns before and after introduction of the incentives yielded several findings (Table II). First, consistent with the analyses of ethnicity and gender seen in Fig. 1, considerable age-related homophily is apparent before the incentive was introduced (see Table II). Older IDUs recruited mostly older IDUs (83.9%), and younger IDUs recruited mostly younger IDUs (65.5%). Homophily is also apparent after the incentives were introduced. However, the degree of homophily changed from a very substantial .49 before the introduction of incentives to .23 afterward. Second, consistent with the expectation that steering incentives would reduce the homophily of groups for whom no steering incentives were offered, the homophily of older IDUs declined, from self-recruitment of 83.9% before the steering incentives to 69.9% afterwards, and so the recruitment of younger IDUs by older IDUs nearly doubled from 16.1 to 30.1%. However, the homophily of younger IDUs also declined, from self-recruitments of 65.5 to 53.1%. This finding is not consistent with the expectation that steering incentives would increase the homophily of groups for whom steering incentives were offered. Third, introduction of incentives was associated with changes in recruitment by age group. Before incentives 42% (19/45) of younger IDUs were recruited by other younger IDUs, this increased to 48% (34/71) after incentives. Therefore, despite their reduced homophily, younger IDUs became more important as recruiters of other younger IDUs. This occurred because after incentives they were more energetic recruiters. While before incentives the 24% of younger IDUs did only 15% (29/190) of the recruiting, after incentives the 39% of them did 34% (64/187) of the recruiting. Thus, the greater postincentive recruitment efforts of younger IDUs and their higher baseline homophily more than compensated for their slightly reduced postincentive homophily. This shows that steering incentives can affect recruitment of a group by altering homophily and by altering the relative success with which different groups recruit.

Assessing Bias in RDS Using Institutional Data

One means for assessing representativeness in RDS is to compare the composition of the sample with institutional data. However, institutional data must be used carefully to assess the representativeness of a sample of a hidden population, for their biases are well known. To assess representativeness of RDS and diminish the geographic filtering effect, we combined the pre- and postincentive samples ($n = 386$).

The number of heroin-related overdoses reported by local hospitals was too small to permit meaningful comparison (Table III). The number of HIV tests for which injection drug use was an admitted risk factor was larger but still modest, and its validity was reduced by the fact that some persons who were tested may not have disclosed injection drug use as an HIV-risk factor. Despite their limitations, the results using these two sources of institutional data were consistent: the largest percentage was for non-Hispanic Whites, followed by Hispanics, with smaller

Table III. Respondents Compared With Institutional Data

		Institutional data				
Percent	Pre- and postincentive respondents 3/95–6/98	Heroin overdoses @ hospital 1/1/96–8/1/98	HIV tests for IDUs 1/1/97–5/30/98	IDUs identified by police 7/1/97–3/15/98	Total population[a]	Population living below poverty line[b]
N	386	10	51	86	59,479	4,342
Non-Hispanic Black(%)	11	0	2	4	3.6	9
Non-Hispanic White(%)	50	90	65	48	82.8	52
Hispanic(%)	38	10	33	48	12.9	39
Other(%)	1	0	2	0	0.7	0
Age 18–25 years(%)	31	40	18	33	—	—
Female(%)	28	30	31	20	—	—

[a] Source, 1990 Census.
[b] Definition of poverty—family of four earning below $12,674 for 1989, source, 1990 Census means data not available.

percentage of cases in the non-Hispanic Black and other categories, and these ordinal rankings are consistent with the RDS data.

Police records on contacts with IDUs gathered during a 6-month period provided a further basis for comparison (Table III). Patrol and other officers were given a data form with which to report contacts with IDUs, where the police definition of an IDU included arrest for possession of drugs or injection paraphernalia, acknowledging injection drug use, or associating with IDUs. When compared to the RDS data, the police data show an overrepresentation of Hispanics, and an underrepresentation of non-Hispanic Blacks. The percentages of non-Hispanic Whites are quite similar. However, the overall ordinal rankings were the same.

The police official who collected and compiled the data forms stated that the 86 IDUs identified by the police constituted in his judgment almost all of the IDUs in Meriden. However, we interviewed 386 IDUs in Meriden. We employed data on the gender, ethnicity, and exact birth date for each of the IDUs identified by police to compute the overlap between the two samples, a procedure that identified 32 respondents. By capture–recapture methods (Wittes and Sidel, 1968), the IDU population can therefore be estimated as $(86/32) \times 386 = 1{,}037$, a figure 11 times larger than the police estimate of the number of IDUs. However, given that our sample was drawn over a longer period, and some mobility into and out of town occurred during the study period, this estimate of 1,037 should only be taken as a rough approximation.

When we compared the racial and ethnic composition of the sample with census data, the discrepancy was large. For example, the sample comprised three times the proportion of Hispanics and non-Hispanic Blacks in the total population. However, given that 74% of our respondents had incomes below the poverty level, a more appropriate comparison was with *impoverished* Meriden residents. The proportions in the census data of non-Hispanic Blacks, non-Hispanic Whites, and Hispanics living below the poverty line in Meriden and the corresponding proportions in the sample differed by less than 2% for each of the four groups of race and ethnicity. Thus, on the basis of this source of institutional data, the sample appears representative.

Assessing Bias Due to Volunteerism and Masking Using Network Data

Despite the hidden nature of the IDU population, a direct test is possible for two sources of bias that have received much attention in the literature on chain-referral sampling: volunteerism (i.e., oversampling cooperative respondents) and masking (i.e., undersampling reclusive respondents). This test involves comparing recruitment patterns with respondents' self-reported personal networks. To assess the size of personal networks, respondents were asked, "How many people do you know by name or street name who also shoot up? These could include sexual partners, running buddies, or other people that you know by first name." Respondents were also asked about the sociodemographic characteristics of these people, including gender, race/ethnicity, age, and homelessness.

In the absence of volunteerism and masking, recruits would be drawn randomly from respondents' personal networks. Therefore, comparing recruitments with these networks provides a means of assessing these sources of bias. This comparison is stratified by age to provide a further assessment of the effect of the steering incentives. For younger IDUs (Table IV)

Table IV. Postincentive Recruits From Meriden Respondents Compared With Members of Personal Networks

	Respondents aged 18–25 years		Respondents aged 26 years or more	
Percent	Recruits ($n = 71$)	IDUs known by name (mean network size = 16)	Recruits ($n = 124$)	IDUs known by name (mean network size = 21.8)
Young (%)	58	51	32	19
Female (%)	32	31	22	33
Non-Hispanic White (%)	51	31	36	32
Non-Hispanic Black (%)	6	7	11	17
Hispanic (%)	44	61	54	51
Other Ethnicity (%)	0	2	0	0
Homeless (%)	19	19	27	17

Note. Only poststeering incentive records are reported, because information on the composition of the personal networks was not available before the change in the incentive system. For example, younger IDUs recruited 58% younger IDUs, but of the IDUs they knew by name, they reported that 51% were younger IDUs.

for gender and homelessness, the composition of recruits and the composition of the personal network differs by no more than 1%. Over recruitment of younger IDUs is 14% (58/51), a finding consistent with the steering incentives. For race/ethnicity, the difference averages 10% across the four groups, reflecting an overrecruitment of non-Hispanic Whites and underrecruitment of Hispanics relative to what would occur if recruitment reflected the composition of the personal networks. Overall, when we excluded age because of the intended over recruitment of younger IDUs, the mean discrepancy between the recruits and the recruiter's personal network averaged 3.7% for the three remaining variables, gender, race/ethnicity, and homelessness. Therefore, aside from the intended effects of the steering incentives, recruitment patterns by younger IDUs approximated their personal network composition.

The association between recruitments and the composition of the personal networks of older IDUs is reported Table IV. Among older IDUs, the greatest proportional difference between recruitment and the composition of the personal networks is the 68% (32%/19%) overrecruitment of younger IDUs, a result consistent with the results of the steering incentives. There is also overrecruitment of the homeless, and underrecruitment of female IDUs. For ethnicity, the differences between recruitment and the composition of the personal networks average less than 4% across the four groups of race and ethnicity. Overall, when we excluded age because of the intended overrecruitment of younger IDUs, the mean discrepancy between the recruits and the recruiter's personal network averages 8.1% for the three remaining variables, gender, race/ethnicity, and homelessness.

The comparison of recruitment patterns and personal networks shows an association between recruitment pattern and personal network composition. When the seven categories of recruitment (i.e., Table IV) are compared with the corresponding personal network composition, the correlation is $p = .874$ for the younger IDUs, and $p = .877$ for the older IDUs, results that also indicate that the association between recruitment and personal network composition is substantial. Therefore, respondents recruited as though they were drawing randomly from their personal networks. These results suggest that masking and volunteerism effects are not strong, and that RDS can potentially serve as a means for reducing these sources of bias in chain-referral samples.

CONCLUSION

The results of this study indicate that steering incentives provide a reasonable strategy for increasing recruitment of a specific population. A further advantage of this method is cost. Recruitment was carried out by those who possess the best information about the population of injectors, that is, other injectors. Furthermore, unlike ethnographers or outreach workers, peer recruiters need not cultivate the trusting relationship required to access injectors for recruitment, because overwhelmingly, respondents recruited those with whom they already had a trusting relationship. When respondents were asked to describe their relationship to their recruiter, this person was described as an acquaintance or friend 90% of the time, and as a relative, sex partner or spouse 7% of the time. The recruiter was described as a stranger only 1% of the time, and the relationship was described as "other" in an additional 2% of cases. Therefore, recruitment overwhelmingly occurred among persons with preexisting relationships. Thus the recruitment process coopted existing social networks among injectors, using those networks as a basis for sampling the target population. Only modest incentives were required to motivate peer recruitment; the cost averaged only about \$14 per recruit. In contrast, in a previous study, when recruitment was carried out by outreach workers in a demographically comparable site, the cost was higher by more than an order of magnitude (Broadhead *et al.*, 1995).

A number of limitations of this study were discussed above, including the use of an interrupted time series design. Future research efforts are needed to validate the results, which suggest that RDS produces samples that are representative of the hidden population from which they were drawn, that bias because of volunteerism and masking is modest, and that steering incentives can be used to increase recruitment of a group of special interest. Improvements in research design could include a pre–post design with a comparison group. Another limitation concerns the reliability of the data with which bias was assessed. Further studies could both apply this sampling method to a nonhidden population, so bias could be directly and definitively assessed; and in further applications to hidden populations, more reliable indicators could be sought.

Future studies should also examine use of steering incentives to increase recruitment of other hidden populations such as the female sexual partners of IDUs and hidden populations other than IDUs.

The prerequisite for any such application is the presence of a *contact pattern* among members of the population; members of the population must engage in some form of activity that brings them into contact with one another. No such contact pattern links persons who engage in spousal or child abuse, and so chain-referral methods are not suitable for studying this hidden population. In contrast, contact patterns link those engaging in sexual or drug-sharing behaviors that place them at risk of HIV infection. Similarly, homeless runaway youth might prove to be a suitable population for study using RDS because they congregate in areas accessible to the homeless and thereby make contact with one another.

APPENDIX: COMPUTING EQUILIBRIUM COMPOSITION OF THE SAMPLES AND THE NUMBER OF WAVES REQUIRED TO APPROXIMATE EQUILIBRIUM IN RESPONDENT-DRIVEN SAMPLES

Computing the Composition of the Equilibrium Sample

Respondent-driven samples attain a stable composition after a modest number of recruitment waves, termed the equilibrium sample composition. Computing this equilibrium requires solving a system of N linear equations, where N is the number of groups into which respondents are divided (Heckathorn, 1997). Where respondents are divided into groups $1, 2, \ldots, n$; S_{xy} is the probability of a member of group X selecting for recruitment a member of group Y, and E_x is the proportion of members of group X in the equilibrium sample $E = (E_1, E_2, \ldots, E_n)$, the system of linear equations is:

$$1 = E_1 + E_2 + \cdots + E_n$$
$$E_1 = S_{11} E_1 + S_{21} E_2 + \cdots + S_{n1} E_n$$
$$E_2 = S_{12} E_1 + S_{22} E_2 + \cdots + S_{n2} E_n$$
$$\vdots$$
$$E_{n-1} = S_{1\,n-1} E_1 + S_{2\,n-1} E_2 + \cdots + S_{n\,n-2} E_n$$

For example, the composition of the equilibrium sample after the introduction of steering incentives can be computed from Table II: $N = 2$, because respondents are divided into two groups, older (1) and younger (2); $S_{11} = .699$, because the proportion of older IDUs recruited by older IDUs is .699 (86/123); similarly, $S_{12} = .301$ (37/123); $S_{21} = .469$ (30/64), and $S_{22} = .531$ (34/64). Therefore, the composition of the equilibrium sample can be computed by solving the following equations:

$$1 = E_1 + E_2$$
$$E_1 = S_{11} E_1 + S_{21} E_2$$

Substituting .699 for S_{11} and .469 for S_{21} yields

$$1 = E_1 + E_2$$
$$E_1 = .699 E_1 + .469 E_2$$

By algebraic manipulation, this reduces to

$$E_1 = .469/(1 - .699 + .469) = .609$$

Finally, given that the equilibrium proportions must sum to 1 (i.e., $1 = E_1 + E_2$), the equilibrium proportion of members of Group 2 is 1 less the proportion in Group 1, (i.e., $E_2 = 1 - E_1 = 1 - .609 = .391$). Thus, the equilibrium composition of the sample is

$$E = (.609, .391).$$

This procedure becomes computationally more demanding as the number of groups increases; dividing respondents into eight groups would require solving a system of eight equations with eight unknowns. Whatever the number of groups, the computational procedures remain the same.

Computing the Number of Recruitment Waves Required to Approximate Equilibrium

To compute the number of waves required to approximate the composition of the equilibrium sample, it is necessary first to identify the way in which the composition of the sample changes from wave to wave. Where X_a^b is the proportion of Group a recruited during Wave b, the proportional distribution of the n groups of respondents during any Wave i is defined by the vector, $X^i = (X_1^i, X_2^i, \ldots, X_n^i)$. The composition of the sample during the subsequent wave, $i + 1$, can be computed as follows:

$$X_1^{i+1} = S_{11} X_1^i + S_{21} X_2^i + \cdots + S_{n1} X_n^i$$
$$X_2^{i+1} = S_{12} X_1^i + S_{22} X_2^i + \cdots + S_{n2} X_n^i$$
$$X_3^{i+1} = S_{13} X_1^i + S_{23} X_2^i + \cdots + S_{n3} X_n^i$$
$$\vdots$$
$$X_n^{i+1} = S_{1n} X_1^i + S_{2n} X_2^i + \cdots + S_{n1} X_n^i$$

Computations begin with wave 0, X^0, which specifies the seeds from which sampling began. Later waves are computed by using the preceeding expressions. After the composition of the wave has been computed, the composition of the wave is compared to the composition of the equilibrium sample. We consider equilibrium to have been approximated when the discrepancy is less than 2% between the equilibrium and wave-specific composition of the sample for each of the sample's n constituent groups.

Respondent-Driven Sampling Software

The procedures for calculating the sampling equilibrium and the number of waves required to approximate the equilibrium are computationally demanding. To facilitate these computations, custom software was developed as part of this project for systems with 20 or fewer groups. For a free copy of this software, please write or e-mail the first author.

ACKNOWLEDGMENTS

This research was made possible by grants from the Division of HIV/AIDS Prevention of the Centers for Disease Control (U62/CCU114816-01) and the National Institute on Drug Abuse (RO1 DA08014). We thank Yaël van Hulst and Robert Mills for their helpful comments and advice.

REFERENCES

Braunstein, M. S. (1993). Sampling a hidden population of noninstitutionalized drug users. *AIDS Education and Prevention*, *59,* 131–139.

Broadhead, R. S., and Fox, K. J. (1993). Occupational health risks of harm reduction work: Combating AIDS among injection drug users. In G. G. Albrecht and R. Zimmerman (Eds.), *Advances in medical sociology: The social and behavioral aspects of AIDS* (pp. 123–142). Connecticut: JAI Press.

Broadhead, R. S., and Heckathorn, D. D. (1994). AIDS prevention outreach among injection drug users: Agency problems and new approaches. *Social Problems*, *41,* 473–495.

Broadhead, R. S., Heckathorn, D. D., Grund, J.-P. C., Stern, L. S., and Anthony, D. L. (1995). Drug users versus outreach workers in combating AIDS: Preliminary results of a peer-driven intervention. *Journal of Drug Issues*, *25,* 531–564.

Broadhead, R. S., Heckathorn, D. D., Weakliem, D. L., Anthony, D. L., Madray, H., Mills, R. J., *et al.* (1998). Harnessing peer networks as an instrument for AIDS prevention: Results from a peer-driven intervention. *Public Health Reports*, *113*(Suppl. 1), 42–57.

Carley, K. (1999). *AIDS Epidemiology Unit Surveillance Report.* State of Connecticut Department of Public Health.

Centers for Disease Control and Prevention (CDC). (1997). *HIV/AIDS Surveillance Report*, *9,* 8–12.

Coyle, S. L., Needle, R. H., and Normand, J. (1998). Outreach-based HIV prevention for injecting drug use: A review of published outcome data. *Public Health Reports*, *113*(Suppl. 1), 19–30.

Diagnosis and reporting of HIV and AIDS in states with integrated HIV and AIDS surveillance—United States, January 1994–June 1997. (1998). *MMWR 47,* 309–315.

Erickson, B. H. (1979). Some problems of inference from chain data. *Sociological Methodology*, *10,* 276–302.

Frank, O., and Snijders, T. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, *10,* 53–67.

Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, *44,* 174–199.

Heckathorn, D. D., Broadhead, R. S., Anthony, D. L., and Weakliem, D. L. (1999). AIDS and social networks: Prevention through network mobilization. *Sociological Focus*, *32,* 159–179.

Heckathorn, D. D., Broadhead, R. S., and Sergeyev, B. (2001). A methodology for reducing respondent duplication and impersonation in samples of hidden populations. *Journal of Drug Issues*, *31,* 543–564.

Holmberg, S. D. (1996). The estimated prevalence and incidence of HIV in 96 large US metropolitan areas. *American Journal of Public Health*, *86,* 642–654.

Johnson, J., Williams, M. L., and Kotarba, J. A. (1990). Proactive and reactive strategies for delivering community-based HIV prevention services: An ethnographic analysis. *AIDS Education and Prevention*, *2,* 191–200.

Klovdahl, A. S. (1989). Urban social networks: Some methodological problems and possibilities. In M. Kochen (Ed.), *The small world* (pp. 176–210). New Jersey: Norwood.

Liebman, J., and Thomas, C. (1992). Development of a sampling strategy to monitor trends in HIV risk behavior in out-of-treatment drug users. In *Epidemiologic trends in drug abuse: Proceedings of the June 1992 Community Epidemiology Work Group* (DHHS Publication No. ADM 92-1958, pp. 494–501). Maryland: Alcohol, Drug Abuse, and Mental Health Administration.

McPherson, J. M., and Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, *52,* 370–379.

Rivera-Beckman, J. (1992). Community outreach in the time of AIDS: The San Francisco, Chicago and POCAAN models. Unpublished manuscript.

Robles, R. R., Colon, H. M., and Freeman, D. H. (1993). Copping areas as sampling recruitment sites for out-of-treatment crack and injection drug users. *Drugs and Society*, *7,* 91–105.

Semaan, S., Des Jarlais, D., Sogalow, E., Ramirez, G., Sweat, M., Norman, L., and Needle, R. (1998). *Meta-analysis of the impact of HIV risk reduction interventions on safer sex behaviors of drug users.* Paper presented at the 12th World AIDS Conference in Geneva, Switzerland, June 30, 1998.

Spreen, M., and Zwaagstra, R. (1994). Personal network sampling, outdegree analysis and multilevel analysis: Introducing the network concept in studies of hidden populations. *International Sociology*, *9,* 475–491.

Sudman, S., Sirken, M. G., and Cowan, C. D. (1988). Sampling rare and elusive populations. *Science*, *240,* 991–996.

Watters, J. K. (1993). The significance of sampling and understanding hidden populations. *Drugs and Society*, *7,* 13–21.

Watters, J., and Biernecki, P. (1989). Targeted sampling options for the study of hidden populations. *Social Problems*, *36,* 416–430.

Wittes, J., and Sidel, V. W. (1968). A generalization of the simple capture–recapture model with applications to epidemiological research. *Journal of Chronic Disease*, *21,* 287–301.